NASA TECHNICAL TRANSLATION NASA TT F-15,656

DEVELOPMENT OF METHODS FOR SYSTEMATIC CONTROL OF CARRIER LIFETIME IN SILICON SLICES

J.L. Lambert, K.H. Sommer, E. Borchert, and J. Koch

Translation of Entwicklung von Verfahren zur gezielten Einstellung der Trägerlebensdauer in Siliziumscheiben, AEG-Telefunken, Frankfurt (West Germany), Report BMFT-FB-T-73-05, April 1973, 194 pp.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
WASHINGTON, D.C. 20546 JUNE 1974

STANDARD TITLE PAGE

1. Report No. NASA TT F-15,656	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle DEVELOPMENT OF METHODS FOR SYSTEMATIC CONTROL OF CARRIER LIFETIME IN SILICON SLICES		5. Report Date June 1974 6. Performing Organization Code
7. Author(s) J.L. Lambert, K.H. Sommer, E. Borchert, J. Koch		8. Performing Organization Report No. 10. Work Unit No.
9. Performing Organization Name and Address Leo Kanner Associates, P.O. Box 5187, Redwood City, California 94063		11. Contract or Grant No. NASw-2481 13. Type of Report and Period Covered Translation
12. Sponsoring Agency Name and Address NATIONAL AERONAUTICS AND SPACE ADMINISTRATION, WASHINGTON, D.C. 20546		14. Sponsoring Agency Code

15. Supplementary Notes

Translation of Entwicklung von Verfahren zur gezielten Einstellung der Trägerlebensdauer in Siliziumscheiben, AEG-Telefunken, Frankfurt (West Germany), Report BMFT-FB-T-73-05, April 1973, 194 pp.

16. Abstract

This report deals with the control of the carrier lifetime in n-type base regions of gallium diffused powder thyristor and diode structures. The scattering of the lifetime values after lowering the lifetime by gold diffusion is due to different vacancy generation rates in different silicon crystals. A practical example of gold diffusion technology is discussed. The phosphorus gettering mechanism is studied. High lifetimes can also be obtained by liquid phase diffusion from gallium. Hall effect measurements indicate the concentrations of deep impurities before and after the high temperature processes. A new method of lifetime measurement developed for process control is described.

17. Key Words (Selected by Author(s))	18. Distribution Statement
	Unclassified - Unlimited

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	166	

Table of Contents

DEVELOPMENT OF METHODS FOR SYSTEMATIC CONTROL OF CARRIER LIFETIME IN SILICON SLICES

J.L. Lambert, K.H. Sommer, E. Borchert, and J. Koch

1. Introduction

/11*

In semiconductor components such as diodes, transistors, and thyristors, the minority carrier diffusion length L is an important parameter. L is a function of the lifetime τ and the mobility μ of these charge carriers. Experience has shown that τ is more variable than μ. Precise knowledge of the lifetime and the extent to which it can be controlled is consequently of great importance for the diffusion length. This work deals with the lifetime control in silicon components prepared by the diffusion technique. The emphasis is on application to thyristor technology.

Adjusting the lifetime in the various p- and n-regions of pnpn-thyristor structures is one of the most important unsolved problems in the modern technology of semiconductor power elements [1]. Very different requirements are made relative to the magnitude of τ. For most types of thyristors, the hole lifetime τ_p in the n-type base should be controllable to within ±20%. In addition, the τ_p-values must be adjustable in a very wide range for frequency thyristors (in the range from 1-10 kHz) in forced-commutation current rectifiers and for extra-high-voltage thyristors for use in high-voltage d.c. transmission systems (HVDCT). A frequency thyristor must have e.g. τ_p-values of 2-4 µsec, while the required lifetimes for thyristors in HVDCT applications are between 50 and 300 µsec.

The influence of τ on various properties of the components mentioned above has already been discussed several times. Among other things, a very strong influence on the transmission characteristic and the recovery time of thyristors has been found. The transmission characteristic was found to be determined by the ratio W/L, where W is the base thickness. In the high-voltage region, Ref. 2 asserts: The larger L is, the greater the conductivity modulation of the base region, and the smaller /12 the transmission losses. The recovery time is particularly closely linked to the lifetime of the surplus charge carriers. Other thyristor properties such as ignition current, holding current, and reverse voltage are functions of the current amplification factors of the two sub-transistors of the thyristor structure and consequently of the τ-values of their base regions [3]. In addition, it has recently been shown [4] that the temperature at which a component is destroyed because of the

*Numbers in the margin indicate pagination in the foreign text.

localization of the forward current rises with increasing carrier lifetime in the base region.

Studies on the influence of manufacturing processes on lifetime are important here. It was discovered very early that the relatively high initial lifetime of silicon crystals (several hundred microseconds, measured by the photoconductive decay method (PCD)) drops to a few microseconds after the high-temperature treatments required for the indiffusion of doping materials. With the aid of chemical analyses, it was found that a primary cause of the reduction in lifetime is impurities diffusing into the silicon where they act as recombination centers. Such impurities are mainly heavy metals such as gold, copper, or iron [5]. The concentrations producing a marked reduction in lifetime are on the order of several ppb. Such trace impurities can originate in the quartz, the source material for the diffusion, or the carrier gases. On the other hand, the influence of nonmetallic impurities -- such as oxygen or carbon -- and of lattice defects on recombination in silicon is still largely unknown [1].

In order to reduce the carrier lifetime to the required small values, only gold has as yet been used in practice. The resulting problems are discussed in detail in this work. In addition, procedures for better control of τ in a large region /13 are investigated. The studies deal with the hole lifetime τ_p in the n-type base (60 Ωcm) of thyristors. The absence of simple methods for systematically controlling the lifetime impairs the yield in the production of conventional thyristors and also obstructs the development of new types.

In Section 2, experimental methods for determining τ_p are discussed, based on measurements of polarity-reversing and turn-off properties of diodes.

In Section 3, the gold diffusion technique for reducing lifetime is described. The reasons for the limited success of this technique are studied, and methods for more reliable carrier lifetime control are given.

Finally, in Section 4, the general features resulting in reductions of carrier lifetimes in pn-structures are discussed. In addition, the effectiveness of getter methods for raising τ_p-values is explained.

2. Determination of Minority Charge Carrier Lifetime τ in Diode Structures /14

While the PCD method is suitable for lifetime determination in homogeneous material, the lifetime in p-n structures can best be determined from the transient properties of prepared p-n diodes. In contrast to the PCD method, which also involves the majority carrier lifetime, the transient behavior of p-n junctions is essentially determined by the minority carrier lifetime. The comparison between the results obtained by different methods for the lifetime is still an object of debate [6].

Two techniques for determining τ utilize the transient properties of diodes:

1. The drop in voltage is measured as a function of time when the current pulse of a p-n junction loaded in the forward direction is instantaneously turned off. This is the so-called post-injection decay method of Lederhandler and Giacoletto [7].

2. The time during which the current remains constant is measured when a diode is switched from the forward to the reverse direction, the so-called retention time t_s. This method was developed by Kingston [8] and by Lax and Neustadter [9].

Both methods are employed in the present work, but predominantly the second.

2.1. Conditions for Measuring τ from the Transient Properties of Diodes

τ can be determined with particular simplicity from the transient properties of diodes under the following conditions:

1. The diode (Fig. 1) must be designed so that the junction poled in the forward direction has an emitter efficiency of about unity. The forward current in the n-region is then carried in essence by holes. /15

2. Weak charge carrier injection is assumed.

3. As in the theories of [7, 8, 9], it is assumed that $W_n \gg L_p$ (diffusion length of holes in n-region).

4. The metallic contact to the n-base does not inject.

5. Volume recombination outweighs any surface recombination.

As explained in [9], the influence of the capacitance of the p-n junction can be neglected for the diodes employed here. These conditions are satisfied in our case. We now describe the measuring techniques mentioned in the introduction.

2.2. Determination of τ

2.2.1. Post-Injection Decay Method

The following discussion deals with the above-mentioned p^+n-junction poled in the forward direction, where the other conditions listed in 2.1 should also be satisfied.

In this method, the source for the forward current is switched off instantaneously at the time t = 0. The post-injection voltage V(t) drops from the value V_0 to 0, corresponding to the charge carrier recombination in the base. The diffusion of charge carriers from the space charge region during this period can be neglected. V(t) is then given by

/16

$$V(t) = \frac{kT}{q} \ln \left[1 + \left(e^{qV_0/kT} - 1 \right) e^{-t/\tau_p} \right] \qquad (1)$$

with k = Boltzmann constant, q = elementary charge, T = absolute temperature. Equation (1) is graphed in Fig. 2a for the case of practical interest, V_0 = 0.52 V = 20 (in units of kT/q). The additional charge carrier density Δp should decrease from the original value Δp(0) by an exponential law of the form

$$\Delta p(t) = \Delta p(0)\, e^{-t/\tau_p} \qquad (2)$$

For the case $V_0 \gg kT/q$, and t/τ_p not very large[1], then

$$V(t) \cong V_0 - \frac{kT}{q} \cdot \frac{t}{\tau_p} \qquad (3)$$

---

[1] more precisely $\frac{qV_0}{kT} - \frac{t}{\tau_p} \gg 1$

It follows from Eq. (3) that V(t) initially drops off linearly (see Fig. 2a). This linear drop has also been observed experimentally. Hence, τ_p can be determined from the slope $\Delta V/\Delta t$ in accordance with

$$\tau_p \approx \frac{kT}{q} \cdot \frac{1}{|\Delta V/\Delta t|} \qquad (4)$$

2.2.2. Retention Time Method

As the current I_F flows through the diode in the forward direction, it is reversed. For a certain time t_s, the so-called retention time, a reverse current I_R continues to flow (see Fig. 2b). After that, the current drops to the value of the saturation current I_S. Physically, this means: the end of the retention time is reached when the charge carrier concentrations at the junction have dropped to their equilibrium value, i.e. the voltage has reached the value zero. Under the conditions in 2.1, we then obtain for the lifetime τ_p of the holes [8]:

$$\operatorname{erf}\sqrt{\left(\frac{t_s}{\tau_p}\right)} = \frac{1}{1 + I_R/I_F} \qquad (5)$$ /17

2.3. Discussion of the Conditions

In the following, we discuss the practical realization of the conditions listed in 2.1.

Emitter Efficiency

Figure 1 shows the p^+n-structures employed for the τ_p measurements. Gallium is diffused into 60 Ωcm (fundamental doping $\sim 10^{14}$ cm^{-3}) phosphorus-doped n-silicon. The thickness W_p of the p^+ layer was 90 µm, the gallium surface concentration $5 \cdot 10^{18}$ cm^{-3}, and the concentration gradient at the p-n junction $3 \cdot 10^{17}$ cm^{-4} [10]. In spite of this flat gradient, studies on the current amplification factor for gallium diffused p^+np^+-structures with a base width $W_n \approx 30$ µm showed that the emitter efficiency is at least 0.85. The emitter efficiency is increased by the inhomogeneous doping of the emitter region. The first condition (see 2.1) is thus satisfied.

Injection Level

The condition of weak injection is satisfied by appropriate choice of the current. The hole current density j in the n-base is given by

$$j = \frac{I_F}{F} = -qD_p dp/dx \approx -qD_p \Delta p/L_p \qquad (6)$$

with D_p = 13 $cm^2 sec^{-1}$ (hole diffusion coefficient). The diameter of the current-carrying area is about 8 mm, so that the area $F \approx 0.5$ cm^2. The donor doping is about 10^{14} cm^{-3}, i.e. the case of weak injection is satisfied for $\Delta p << 10^{14}$ cm^{-3}, namely: /18

$$\Delta p = \frac{I_F}{F} \cdot \frac{L_p}{D_p} \cdot \frac{1}{q} << 10^{14} \text{ cm}^{-3}$$

with L_p = 100 µm implies $I_F << 10$ mA.

This requirement is obviously too strict in the present case. As shown in 2.4, there is no appreciable effect for I_F up to 20 mA.

Base Width W_n

If the conditions listed are not satisfied, it will not be the actual lifetime τ_p which is determined by way of (5), but only an "effective lifetime" τ_p^{eff}. Figure 3 shows the influence of W_n on τ_p^{eff}. It can be seen that for $W_n \geq 4.5\ L_p$, τ_p^{eff} is independent of W_n. This is consistent with the results of other authors [11, 12]. Hence, in this work, only diodes with a thick base in the sense described (i.e. $W_n \geq 4.5\ L_p$) are employed for τ_p measurements.

n-Base Contact

If the theories in [7, 8, 9] are valid, no charge carrier injection should result from n-base contact; i.e. the charge carrier concentrations should not deviate from their equilibrium values there during the switching process. This condition is satisfied by an ohmic contact. It can be simply prepared by rubbing the silicon with a gallium-wetted aluminum rod. The rubbing probably disturbs the silicon surface to such an extent that high recombination rates are achieved at the metal-silicon interface. In order to make sure of the ohmic characteristic of such a contact, current-voltage measurements were performed for n^+nm-structures with a diameter of 10 mm. The n^+ layer was formed with an Au-Sb-alloy contact on the n-base; m designates the gallium-aluminum layer on the silicon. The studies extended to currents of 50 mA and to structures with a base resistance /19 between 50 and 100 Ωcm. The measurements actually found an ohmic characteristic for these structures.

Diode Geometry and Influence of Surface Recombination

The diodes employed were prepared from large diffused p^+np^+-slices by ultrasonic drilling and lapping off a p^+-layer. They had a diameter of 10 mm. They were then etched for about 10 sec in CP 6, which reduced the surface recombination rate to values on the order of 10 cm/sec [13]. As depicted in Fig. 1, a metal contact 5 mm in diameter was applied to the p^+-layer with the aid of a gallium-wetted aluminum rod, and a contact 10 mm in diameter was applied to the base layer. Apart from its simplicity, the application of this procedure also has the advantage that no additional high-temperature treatment is required. Before the lifetime measurements, the current-voltage characteristic of the rectifiers was checked; diodes with a reverse current of more than 50 μA at 50 V were discarded.

2.4. Experimental Determination of Minority Carrier Lifetime

The τ_p values were determined for a number of such diodes by both transient methods. The measurements were taken with a Tektronix 545 A oscillograph with a Type B plug-in unit. The forward current I_F could be set to the values 1, 2, 5, 10, and 20 mA and the reverse current to 0.1, 0.2, 0.5, and 1.2 mA. The duration of the forward current pulse was 500 μsec, i.e. long enough to bring about the steady-state minority carrier distribution. The reverse voltage pulse in the Kingston method was 1500 μsec. It was not the current, but the voltage which was measured.

The results of the measurements are reproduced in Table 1. /20 They show that the lifetime values in these diodes were between 2 and 130 μsec. It gives τ_p^{PI} obtained with the post-injection voltage method, the values t_s obtained by Kingston's method, and the values of τ_p^S calculated from (5). I_F ran through the given values from 1 to 20 mA, and 0.1 and 1.0 mA were chosen for I_R in the τ_p^S determination. In the τ_p^{PI} determination, it was observed that the linear part of the descending voltage-time curve for $I_F < 20$ mA became less and less distinct as the current decreased. The τ_p^{PI} values for $I_F = 5$ and 1 mA are therefore only estimated. The mean value $\bar{\tau}_p^S$ of τ_p^S over the different forward currents at the two selected reverse currents and the lifetime value τ_p^{PI} for $I_F = 20$ mA are tabulated for the different diodes in Table 2.

2.5. Discussion of the Results

Except for diode 6, τ_p^S is more or less independent of the injection level and, as expected, also of I_R. The values for τ_p^{PI} are smaller than the calculated τ_p^S values. For the diodes with small lifetime values, τ_p^{PI} is about 50% of the corresponding τ_p^S and reaches about 75% of it as the lifetime increases. As mentioned above, the post-injection voltage drop was very difficult to evaluate for small currents. Experimentally, it was always simpler to determine the t_s values than the slope of the voltage drop. Therefore, in this work, we mainly use the retention method. The important question of which of the two methods yields the more accurate values for τ_p is presently the object of a special study [6].

3. Control of Minority Carrier Lifetime in Thyristor Structures Through Gold Diffusion /21

3.1. Introduction

This chapter deals with the controlled adjustment of gold concentration and thus the lifetime in diffused silicon slices through solid-state diffusion. The gold diffusion process is employed for virtually all types of silicon components, and is thus of great technological importance. However, various aspects of the process are not yet sufficiently understood. Primarily, these are (a) the anomalous diffusion behavior, expressed in nonideal concentration profiles; and (b) the structure-dependence of diffusion, manifested by the scattering of lifetime values after gold diffusion from one crystal to the next. Another poorly understood phenomenon, which is important for component manufacture, is the influence of diffusing n^+ and p^+ layers (with small resistance) on the gold distribution in the slices. The main part of this chapter is devoted to the investigation of these problems.

Specifically, these problems are:

(a) It is known that the gold concentration C in silicon slices cannot be described in a simple manner by Fick's second law:

$$\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial x^2} \qquad (7)$$

In particular, it is not possible to describe the process by a simple diffusion constant D. If gold is allowed to diffuse into a slice from both sides, the observed gold diffusion

profiles are instead more or less independent of position in the middle region. At the ends, C(x,t) rises to high surface concentrations. In Fig. 4, a typical gold diffusion profile is compared with ideal profiles, obtained by solving (7) [37] with the initial condition C(x) = 0 for t = 0 and the boundary condition C/C' = 1 for x = 0, x = 2 (C' is the saturation value of C) for various times. The gold profile was taken from the results of Martin et al. [14], who diffused gold into a slice 300 µm thick for 20 minutes at 720°C. Similar gold profiles were obtained at other diffusion temperatures and for other diffusion times by Sprokel and Fairfield [15] and by Gilpin and Boatman [16]. The causes of these anomalous gold profiles and the practically important question of the time-dependence of gold concentration in the middle region has not yet been adequately investigated. Hence, they are given special attention here. By means of experimental results of the author, it is shown in Section 3.4 that the gold concentration in the middle region can be represented by a root function of the diffusion time. /22

(b) It will turn out that the gold concentration in silicon after a fixed finite diffusion time depends very sensitively on the perfection of the crystal structure. For the production of components, it is important to know this, since the degree of imperfection of the starting material is substantially determined by the crystal-growing procedure employed, which thus indirectly influences the properties of the components. This material-dependence of gold diffusion is particularly important for the production of power components such as thyristors, for which a relatively large volume of silicon is required. In the following, the specific causes of the material-dependence will be investigated and methods to overcome this difficulty described.

3.2. The Irreproducibility of the Previous Gold Diffusion Process

We consider the following case from practice. A 200-A, 1000-V inverter thyristor for middle-frequency applications (∿5 kHz) is to have a recovery time of 30 ± 5 µsec. Empirically, it is found that a hole lifetime of 2.5 ± 0.3 µsec is required, based on the lifetime value from the storage charge. The reasons for these small tolerances in the value of τ_p are discussed in Section 1. In producing the desired $n^+p^+np^+$-structure from n-type silicon, however, τ_p now drops from several hundred microseconds in typical cases to 20-30 µsec, i.e. to a higher value than the desired one of 2.5 µsec. Hence, gold diffusion is also necessary. /23

Previously, it was assumed that a reproducible gold concentration was introduced into a silicon structure of given

dimensions under fixed diffusion conditions. Table 3 shows that this is not the case. At 785°C, gold was diffused for an hour simultaneously into a number of p^+n-diodes of different silicon crystals, but with the same p^+ layer thickness and surface concentration and the same resistivity (50 Ωcm) in the n-type layer. Gold diffusion was carried out for a similar group of diodes for an hour at 805°C. The evaporation of the gold and the prior cleaning of the surface were always subject to the same conditions. As can be seen from the table, the values for τ_p after both diffusions fluctuate considerably, varying by a factor of almost 4.

Nevertheless, the lifetime $\tau_p(t)$ after diffusion is not independent of the lifetime τ_p'' before diffusion. If the recombination centers are independent of each other, we have:

$$\frac{1}{\tau_p(t)} = \frac{1}{\tau_p''} + \frac{1}{\tau_p'(t)} \qquad (8)$$

where τ_p' means the carrier lifetime corresponding to the indiffused gold concentration. The values of τ_p', which are accordingly calculated from τ_p'' and τ_p, also vary widely. However, in many cases, τ_p'' is much larger than τ_p, so that $\tau_p'(t) \approx \tau_p(t)$. In general, it can be inferred from the scattering of $\tau_p'(t)$ that gold diffusion is material-dependent.

To overcome this difficulty, the manufacturer can proceed /24
in the following manner: Let gold diffuse into a certain group of $n^+p^+np^+$ structures. Gold is initially diffused into some test structures of the group under arbitrary conditions, e.g. for an hour at 830°C. From these gold-diffused structures, p^+n-diodes are now produced, with the aid of which $\tau_p(t)$ is measured. If $\tau_p(t)$ is larger than the desired value, other samples are gold-diffused at a higher temperature and vice versa. This is repeated until the desired lifetime is achieved with the required accuracy, at which point gold diffusion is carried out for the entire group. Since, in our experience, as many as five test diffusions are required, this procedure is time-consuming and unsatisfactory.

3.3. Mechanism of Gold Diffusion in Silicon

A better solution to the problem described in Section 3.2 requires more precise comprehension of the mechanism of gold diffusion in silicon. In this section, we will therefore give a brief description of the previous work in this area, including theoretical and experimental contributions of the author as well.

In previous works, conflicting values were found for the diffusion constant $D = D_0\, e^{-E/kT}$ of gold in silicon; Struthers

[17] found $D_0 = 0.011$ cm^2sec^{-1} and $E = 1.32$ eV, while Boltaks [18] obtained $D_0 = 2.44 \cdot 10^4$ cm^2sec^{-1} and $E = 3.11$ eV. Both authors investigated the temperature range 800° to 1200°C and obtained their results by fitting the solutions of (7) to the profiles found, i.e. under the assumption of an ideal diffusion process. As was reported by Wilcox and La Chapelle [22], the gold profiles assume a virtually ideal form for relatively long diffusion times in "thick" samples (5 mm); from this form an "effective" diffusion constant can be determined. The different results of Struthers on the one hand and Boltaks on the other might therefore be a manifestation of such a material- /25
dependent effective diffusion constant.

Dash [19] obtained indirect evidence for the nonideal behavior of gold diffusion. He was primarily interested in the motion of dislocations in silicon during gold diffusion. By careful grooving, he generated a small number of left-oriented screw dislocations, and observed that the latter were converted to right-oriented screw dislocations by subsequent gold diffusion at 1000 to 1300°C. Dash concluded that the screw dislocations had climbed during the gold diffusion. He also deduced that this climbing process had to be connected with generation, but not annihilation, of lattice vacancies. Namely, it could be shown that the extent of climbing was related to the diffusion time and was independent of the cooling rate after diffusion. To explain this result, Dash proposed that gold diffusion in silicon proceeded by a dissociative substitution-interlattice mechanism, previously introduced by Frank and Turnbull [20] to explain the anomalous rapid diffusion of copper in germanium. In accordance with the Frank-Turnbull mechanism, the diffusing atoms can move both in the interlattice as well as through lattice sites, where interstitial diffusion proceeds much more rapidly than lattice diffusion. The diffusing atoms then have certain interlattice and lattice-site solubilities. The diffusion current is now carried more or less by interstitial atoms alone, a few of which now fall into vacancies and are thus substitutionally incorporated. The climbing of dislocations during the gold diffusion was attributed by Dash to the generation of vacancies.

3.3.1. The Frank-Turnbull Mechanism of Gold Diffusion in Silicon /26

The principle of the Frank-Turnbull mechanism, as developed by these authors and later by Sturge [21] for copper in germanium, is now employed for the case of gold diffusion in silicon. An essential feature of this mechanism is the conversion of rapidly diffusing interstitial atoms (Au_I) into substitutionally incorporated atoms (Au_S) by trapping in vacancies (O) in accordance with the relationship

$$Au_S \rightleftarrows Au_I + O$$

The reaction velocity is then given by

$$\tau^o \frac{\partial C_S}{\partial t} = K C_I C_V - C_S \tag{9}$$

where C_S and C_I are the concentrations of the gold atoms in lattice and interlattice sites respectively and C_V is the concentration of vacancies. For the constant K, we have:

$$K = \frac{C'_S}{C'_I C'_V}$$

The apostrophes on the concentrations indicate the corresponding equilibrium values. These values are assumed to depend only on temperature.

The diffusion of vacancies and that of gold atoms over interlattice and lattice sites is described by the equations

$$\frac{\partial C_V}{\partial t} = D_V \frac{\partial^2 C_V}{\partial x^2} + g_V + \left.\frac{\partial C_I}{\partial t}\right|_S \tag{10}$$

$$\frac{\partial C_I}{\partial t} = D_I \frac{\partial^2 C_I}{\partial x^2} + \left.\frac{\partial C_I}{\partial t}\right|_S \tag{11}$$

$$\frac{\partial C_S}{\partial t} = D_S \frac{\partial^2 C_S}{\partial x^2} + \left.\frac{\partial C_S}{\partial t}\right|_I \tag{12}$$

/27

Here, g_V is the vacancy generation rate and

$$\left.\frac{\partial C_I}{\partial t}\right|_S \quad \text{and} \quad \left.\frac{\partial C_S}{\partial t}\right|_I$$

the time derivatives of C_I and C_S respectively, to the extent that they are determined by the lattice site ⇄ interlattice site transition. Naturally,

$$\left.\frac{\partial C_I}{\partial t}\right|_S = -\left.\frac{\partial C_S}{\partial t}\right|_I \tag{13}$$

From (9) with the condition $\tau^{o} = 0$, we obtain the relationship

$$C_S = K\, C_I\, C_V \qquad (14)$$

3.3.2. Relative Solubilities and Diffusion Constants

In the following, the relative values of D_S, D_V, D_I and C_S^i, C_V^i, C_I^i will be discussed as functions of temperature. These values, given by different authors -- in particular by Wilcox and La Chapelle [22] -- are depicted in Figs. 5 and 6. Although the values of D_S and D_I are only estimates, they still ought to be of the same order of magnitude as the true values. It is well known that atoms such as phosphorus and gallium diffuse primarily substitutionally and thus have diffusion constants of the order of magnitude of the D_S values shown in Fig. 5. Analogously, atoms such as copper and iron which prefer to diffuse interstitially have diffusion constants of the same order of magnitude as D_I. The values for D_V and C_V^i were taken from the literature for double vacancies, since it is unlikely that simple vacancies participate to a significant degree in the diffusion process in silicon [23]. In Fig. 6, the values for C_S^i/C_I^i were taken from [22]. The values for C_S^i/C_V^i were calculated from values given by Bullis [24]. The sources for the values of $C_S^i(T)$ in Fig. 7 and of $C_V^i(T)$ (double vacancies) /28 are Bullis [24] and Kendall and De Vries [23] respectively.

From Figs. 5 and 6 it can be seen that in the temperature range of interest (700 to 1000°C), the following relations hold:

$$D_S \ll D_V \ll D_I, \qquad C_S^i \gg C_V^i, \qquad C_S^i \gg C_I^i$$

In the following study of gold diffusion in silicon, we deal with the case of diffusion in thin (500 to 800 µm) slices, which is of interest from the practical standpoint. In view of the relatively large value of $D_I = 5 \cdot 10^{-6}$ cm^{2}sec^{-1} at 900°C, the maximum interstitial solubility for interlattice atoms will be achieved after times on the order of some 10 sec. Thus, for example, the diffusion length $\sqrt{(D_I t)} = 500$ µm at 900°C for $t = 50$ sec. The assumption $C_I = C_I^i$ is therefore sure to be satisfied when considering diffusions in thin slices for a period of more than 1 hour.

3.3.3. Diffusion Model

The following model is formulated for the sequence of events in the diffusion process:

At the beginning of the diffusion process, because of $D_I >> D_V >> D_S$, atoms diffuse homogeneously within 1 minute up to the saturation concentration C_I^+. Then $\partial C_I/\partial t \approx 0$. Because of the high diffusion constant D_I of the first term on the right side of (11), transitions from interstitial sites to lattice sites are virtually completely "intercepted" by instantaneous replenishment. In accordance with (14), equilibrium is achieved between C_S, C_I, and C_V. The equilibrium point at the end of this phase will be required for the following calculation and will therefore be estimated in terms of order of magnitude: /29

Addition of (10) and (12), using (13), and neglecting diffusion yields the equation

$$\frac{d(C_V + C_S)}{dt} = g_V$$

If there were no generation of vacancies, we would have

$$C_V(t) + C_S(t) = C_V^+$$

because of the initial conditions $C_S(0) = 0$ and $C_V(0) = C_V^+$. Under this condition, we would then obtain with (14) the relationship

$$C_V(t) = \frac{C_V^+}{1 + K\, C_I(t)} \tag{15}$$

This value is a minimum for C_V, since generation produces a value larger than C_V^+. At the end of this phase, we obtain from (15), because of $C_I = C_I^+$, in orders of magnitude

$$\frac{C_V}{C_V^+} = \frac{C_V^+}{C_S^+ + C_V^+} = \Theta \ll 1, \tag{16}$$

where K is expressed in terms of the saturation values in accordance with (14). It will be found that the precise value of Θ does not enter into the final result which is our objective, as long as this last inequality is satisfied. That is, as long as the vacancies $\int g_V\, dt$ produced during this phase by generation do not change anything in this inequality, (16) can be employed as an initial condition for the subsequent phase. In this second phase which now follows, we can set $C_I = C_I^+$, but now the diffusion of vacancies and gold atoms to lattice sites must be taken into account. Adding (10) and (12), and using (14), we obtain

$$\frac{\partial C_S}{\partial t} = D' \frac{\partial^2 C_S}{\partial x^2} + \frac{C_S'}{C_V' + C_S'} g_V; \quad D' = \frac{C_S' D_S + C_V' D_V}{C_S' + C_V'} \tag{17}$$

D' was depicted as a function of temperature in Fig. 5.

A simple expression for g_V is obtained from the mass action law as follows: Assuming a constant generation rate K' and a recombination rate proportional to the existing vacancy concentration, i.e. $K'C_V$, we obtain for the total generation /30

$$g_V = K' - K''C_V \tag{18}$$

In thermal equilibrium, $g_V = 0$, $C_V = C_V'$, i.e.

$$0 = K' - K''C_V' \tag{19}$$

Substituting (19) in (18) yields

$$g_V = A\,(1 - C_V/C_V')\,, \tag{20}$$

where $A = K'C_V'$.

With (17), (20) and the relationship resulting from (14):

$$\frac{C_V}{C_V'} = \frac{C_S}{C_S'}$$

we obtain

$$\frac{\partial C_S}{\partial t} = D' \frac{\partial^2 C_S}{\partial x^2} + A\,(1 - \frac{C_S}{C_S'})\,. \tag{21}$$

Partial results can be estimated in a simple fashion from (21): The initial condition following from (16)

$$\frac{C_S(x,0)}{C_S'} = \Theta \tag{22}$$

states that in practice there is a distribution over lattice sites which is far under the equilibrium concentration. In accordance with (21), this concentration can either be supplemented by diffusion in from the edges, or by generation (second term on the right side). As long as diffusion has not yet

reached the middle region, only the generation term on the right side of (21) will make a perceptible contribution to increasing the concentration. Therefore, in the middle region, the change in concentration is correctly described by /31

$$\frac{dC_S}{dt} = A\left(1 - \frac{C_S}{C_S'}\right)$$

with the initial condition (22), for reasonably short times. With $k' = A/C_S'$, the solution then reads

$$\frac{C_S}{C_S'} = 1 - (1 - \theta)\, e^{-k't}. \qquad (23)$$

It can be seen that for times which are not too short, the solution for $\theta << 1$ is practically independent of θ. As will follow from the experimental study of gold diffusion to be described in the next Section 3.4, (23) does not supply the right time-dependence for C_S/C_S'. The experimental results show that C_S/C_S' is proportional to $\sqrt{t}$ for $k't << 1$ and does not vary linearly with t, as predicted by (23). Nevertheless, this model permits a qualitative interpretation of the homogeneous, increasing (in time) gold concentration in the middle region:

At the beginning, because of the high diffusion constant D_I, only gold atoms diffuse over interlattice sites and distribute themselves homogeneously over the thin slice with the relatively low equilibrium concentration C_I'. The existing vacancies will be largely filled by gold atoms. However, the concentration will be far below the equilibrium density. Further gold atoms will not be able to transfer from interlattice sites to lattice sites until the vacancy concentration is raised at the position involved. Because of the large diffusion constant D_I, the replenishment of gold atoms via interstitial sites takes place almost "instantaneously." Because of the relatively low diffusion constant D_V, vacancies can be replenished by diffusion practically only in the boundary regions, while in the middle region, the replenishment of vacancies can only be accomplished by generation.

There is a further discussion of the diffusion and generation mechanism in Section 3.4.4.

3.4. Experimental Investigations of the Gold Diffusion Process /32

In this section, experimental investigations of the time-dependence of gold concentration in the flat segment of the

profile in silicon slices by means of neutron activation analysis are described. The determination of gold in silicon by activation analysis is simple and still very precise, even with very small quantities to be analyzed ($\approx 10^{-12}$ g). This high sensitivity is achieved because of the large effective cross section for the capture of thermal neutrons and the favorable half-life of the resulting gold isotope Au 198. Thus, for example, a gold concentration of 10^{10} cm^{-3} can be determined in a sample of silicon weighing 1 g.

3.4.1. Experimental Methods

Slices about 800 µm thick are cut out of cylindrical single crystals about 1 cm in diameter. These slices are carefully cleaned in an ultrasonic bath, degreased, and then etched in a CP6 etching solution (2 parts HNO_3, 1 part CH_3COOH, 1 part HF); a gold layer 0.5 µm thick is deposited from the vapor onto the etched surfaces. The slices prepared in this way were placed in a three-zone diffusion furnace. The inert gas argon was circulated through at a rate of 60 ℓ/min. Its temperature could be held constant within ±1°C over a length of 15 cm. The gold was indiffused at a fixed diffusion temperature between 750 and 1100°C. The diffusion time was between 1 and 100 hours. Then, by withdrawing them quickly from the furnace, the specimens were cooled to room temperature at a rate of about 10°C/sec. The non-indiffused gold was easily removed by an etching solution of HNO_3 and HCl. A silicon layer 150 µm thick was removed on both sides, the first 100 µm mechanically by lapping, and the remaining 50 µm chemically by means of CP6. To determine the diffused-in gold radiochemically, the slices were then activated for 3 days in the Karlsruhe nuclear reactor with a neutron flux density of 10^{13} $cm^{-2}sec^{-1}$. After the silicon activity had died away (i.e. after roughly 24 hours), another 100 µm silicon was removed chemically from both sides of the slices, in order to eliminate any possible surface contamination with gold resulting from incautious handling of the specimens between the last etching step and the end of the irradiation. Accordingly, of the 800-µm-thick slice, only a thin 300-µm plate remained for the study, i.e. only the middle of the original specimen. This also makes sure that the measured gold concentration will not be artificially elevated by including the original surface regions because of the possibly higher gold concentrations there (cf. Fig. 4). /33

The determination of the gold distribution in the specimen was carried out with the aid of known radiochemical techniques.

3.4.2. Gold Concentration as a Function of Time in Middle of Silicon Slice with High Dislocation Density

The first studies were carried out on silicon slices containing about 10^4 dislocations/cm^2 (crystal A). The dislocation density was determined by etch pit counts. The source of the material was Wacker Co., Munich. The material was zone-pulled and n-conducting with values for the resistivity ρ between 40 and 60 Ωcm.

These slices were gold-diffused at 835°C over various times between 1 and 100 hours, each slice never being used for more than one diffusion. By means of profile measurements on the sample with the diffusion time of 1 hour, it was ensured that a flat profile could be produced in the center of the sample under these experimental conditions as well. (For this purpose, with the aid of a slow-acting etchant, a layer about 30 µm thick was removed on both sides of the initially 300-µm slice.)

The results of the gold-concentration measurements in the middle of the specimen are plotted on a log-log graph in Fig. 8 (curve A). The gold concentration is expressed in terms of the solubility C_S^i (at 835°C, $C_S^i = 8.0 \cdot 10^{14}$ cm^{-3} [24]). Figure 8 shows that the curve obtained can be described well by the relationship (cf. (23)) /34

$$\frac{C_S}{C_S^i} = \sqrt{\left(1 - e^{-k_o t}\right)} \qquad (24)$$

For $k_o t \ll 1$, we obtain the square-root function

$$\frac{C_S}{C_S^i} = \sqrt{k_o t} \qquad (25)$$

From Fig. 8 (curve A), we obtain a value of $k_o = 1.34 \cdot 10^{-2}$/hour.

3.4.3. Gold Concentration as a Function of Time in Middle of Silicon Slice with Low Dislocation Density

Under the same experimental conditions as described in Section 3.4.2 for silicon with 10^4 dislocations/cm^2, C_S/C_S^i was determined as a function of t for "dislocation-poor" silicon as well. Two crystals (B, C) were studied which were prepared under different conditions: crystal B was prepared using a modified zone-melting procedure ("Lopex" material, Texas Instruments Co., Dallas, Texas). Crystal C was obtained

by the Czochralski procedure using a quartz crucible (Wacker Co., Munich). Both crystals were n-conducting with $\rho \simeq 50\ \Omega cm$. In comparison with crystals A and B, which had concentrations of dissolved oxygen less than $10^{17}/cm^3$, crystal C contained more oxygen ($10^{18}/cm^3$).

Etch pit studies again were initially employed to determine the dislocation density. In crystal C, the count of etch pits yielded a surface density of $N \simeq 10^2/cm^2$. The density was higher in crystal B. The interpretation of the pits as etched points /35 at which the dislocations penetrate the surface was checked by x-ray crystallography, and this confirmed only the value $N \simeq \simeq 10^2/cm^2$ for the crystal C. On the other hand, the crystal B appeared free of dislocations in the (more reliable) x-ray diffraction pattern. The etch pits must therefore have arisen at defects of a different nature. Etching effects which have nothing to do with dislocations are known. However, they were not further investigated here.

The measured dependence of C_S/C_S^* on t for the crystals B and C is likewise plotted in Fig. 8. Here too, (24) describes correctly the experimentally determined curve. The constants are $k_o = 1.9 \cdot 10^{-3}$/hour for crystal B and $k_o = 1.4 \cdot 10^{-3}$/hour for crystal C, i.e. much smaller than for crystal A. Allowing for the measuring accuracy of ±5%, with which the gold concentration was determined with the aid of activation analysis, and in spite of the greater scattering of the measurements observed, it still appears that a genuine effect is involved, which can perhaps be attributed to local fluctuations of k_o in the crystal. Further evidence in support of this hypothesis is found later in connection with the discussion of the variation of carrier lifetime with time during gold diffusion.

Additional values for $C_S(t)/C_S^*$ are incorporated in Fig. 8 from the work of Sprokel and Fairfield [15] for Czochralski-pulled material of low dislocation density for two gold-diffusion temperatures (1000, 1100°C). It can be seen that these results are also correctly reproduced by (24), here with $k_o = 5.0 \cdot 10^{-3}$/hour at 1000°C, and $k_o = 4 \cdot 10^{-2}$/hour at 1100°C.

Further measurements of $C_S(t)/C_S^*$ at various temperatures are plotted in Fig. 9 for the dislocation-free crystal B. The resulting values of $k_o(T)$ are depicted in Fig. 10 as a function of 1/T; the energy 2.3 eV can be calculated from the slope of the Arrhenius lines. A corresponding energy of 2.9 eV is obtained from the two values of k_o of Sprokel and Fairfield [15].

Neither the temperature-dependence of k_o nor the experimentally observed $\sqrt{t}$-dependence of gold concentration has yet been successfully explained. The time-dependence can only be understood with the aid of corresponding vacancy generation rates. Heigl and Sizmann [25] found an increase in vacancy concentration proportional to $\sqrt{t}$ in platinum. /36

3.4.4. Vacancy Sources in Silicon

Although (24) has not yet been interpreted theoretically, it is still possible to come up with some ideas on the physical importance of k_o. As can be seen in Fig. 8, crystal A (10^4 dislocations/cm^2) has a k_o which is greater than that of the low-dislocation-density crystals B and C by about a factor of 10. Since dislocations can serve as sources of vacancies, k_o obviously depends on their effective density V_Q. The larger k_o is, the greater V_Q must be and the faster C_S will converge to its saturation value.

Different values for k_o or V_Q in different crystals have great practical importance, because they inevitably result in different gold concentrations in the crystal for identical diffusion times. The irreproducibility of carrier lifetime after gold diffusion found in Section 3.2 should be due to this fact.

However, the results obtained from crystals B and C imply that, apart from the embedded edge dislocations, there must be other lattice defects active as vacancy sources in commercial silicon crystals. John [26] gives a description of lattice defects found (in addition to dislocations) by electron microscope in silicon "with semiconductor character," such as that which we have studied:

1. rather large inclusions of foreign substances, several μm in diameter,
2. segregations rich in silicon, particularly SiO_2 and SiC,
3. ring-shaped defects, probably silicon-vacancy clusters, /37
4. so-called etch-hills, and finally
5. star-shaped formations, which can be associated with a local breakdown of the silicon lattice.

Recently de Kok [27], in etching experiments on single dislocation-free, high-resistance silicon crystals, found other ring-shaped etch-pit patterns, and also interpreted them as

etched vacancy clusters, which could arise in a crystal saturated with vacancies if it were cooled very rapidly. Another phenomenon, the so-called "doping rings" [26], can also be confidently associated with macroscopic lattice defects. These are ring-shaped fluctuations in electrical resistance around the crystallization axis. They have been detected in silicon, independent of its semiconductor quality, and are considered unavoidable in the usual crystal preparation procedures.

3.5. Control of Charge Carrier Lifetime in Gallium-Diffused p^+n-Diodes by Means of Gold Diffusion

After discussing the gold diffusion process in silicon, we now investigate problems associated with precise control of the hole lifetime τ_p in the n-region of a p^+n-diode. Between τ_p and the gold concentration, there is the following relationship

$$\tau_p = \frac{1}{N} \cdot f\left(\alpha_{n,p}, \beta_{n,p}, n, p\right) \tag{27}$$

This excludes other possible recombination centers as indiffused substitutional gold atoms. Here $\alpha_{n,p}$ and $\beta_{n,p}$ are the electron and hole capture cross sections for the gold acceptor or donor levels in silicon, and n and p are the concentrations of free charge carriers, which determine the charge carrier injection level [24]. Assuming $C_S >> C_I$, then $N = C_S$. This means that the substitutional gold is the only significant recombination center. Using (27), and assuming that the diffusion conditions for gold are given by (24), $\tau_p(t)$ can be described after a diffusion time t by the function /38

$$\frac{1}{\tau_p(t)} = \frac{1}{\tau_{min}} \sqrt{\left(1 - e^{-k_o t}\right)} \tag{28}$$

τ_{min} is the minimum saturation value of charge carrier lifetime for $t \to \infty$ and $C_S \to C_S'$.

In order to check (28) experimentally, two important prerequisites must be satisfied:

a) Before the beginning of gold diffusion, no other active recombination centers can be present, but the charge carrier lifetime τ_p'' before diffusion must at least be much larger than the value $\tau_p(t)$ after diffusion, since

$$\frac{1}{\tau_p(t)} = \frac{1}{\tau_p''} + \frac{1}{\tau_p'} \tag{8}$$

The lifetime τ_p' refers only to indiffused gold. Since $\tau_p'' >> \tau_p(t)$, we obtain $\tau_p' = \tau_p(t)$.

As will be described later, $\tau_p'' > 100$ µsec can actually be achieved in the preparation of p^+n-diodes by means of liquid gallium sources. Since the largest measured values of $\tau_p(t)$ after 1-hour gold diffusion were on the order of 10-15 µsec, the condition for $\tau_p' = \tau_p(t)$ was therefore satisfied in our experiment. The corresponding τ_p values were measured by the retention-time method (I_F = 5 mA, I_R = 1 mA).

For small values of τ_p, such as those obtained in conventional vapor-phase gallium diffusion for preparation of p^+n-diodes, $\tau_p'' \approx \tau_p(t)$, and thus $\tau_p(t)$ can be represented by

$$\frac{1}{\tau_p'} = \frac{1}{\tau_p(t)} - \frac{1}{\tau_p''} = \sqrt{(1 - e^{-k_0 t})} \qquad (29)$$

b) If the p^+n-diodes are quenched from the diffusion temperature to room temperature at the end of the diffusion process, additional recombination centers can be generated in addition to the indiffused gold atoms. Silicon is amply deformable above about 1000°C. Mechanical stresses created in a slice if it is rapidly cooled from temperatures above 1000°C relax via slipping processes with the formation of dislocation networks [28]. This slippage causes either directly or indirectly a decrease in charge carrier lifetime, probably due to the creation of dangling bonds. Hence, in order to check (28) experimentally, slipping processes must be avoided. This was achieved by choosing a gold diffusion temperature in the nonplastic range of silicon, namely 850°C. The following experiment proved that this temperature is low enough. A diode was placed in the diffusion furnace at 850°C, and then rapidly withdrawn after thermal equalization (after about 10 sec). The temperature of the slice was measured with an optical pyrometer. Measuring τ_p after the quenching experiment yielded the original value of the lifetime τ_p'' = 80 µsec. If the experiment was repeated at 1050°C with an identical diode from this series, however, a decrease in τ_p from 82 µsec to 19 µsec was observed. However, this decrease in τ_p is ascribed to the quenching process, since under these conditions, no appreciable gold diffusion into the n-base can have taken place. With $D_I = 7 \cdot 10^{-6}$ $cm^2 sec^{-1}$ (1000°C), the diffusion length is $\sqrt{(D_I t)} = 84$ µm for t = 10 sec. The gold concentration was therefore raised essentially only in the p^+-zone. /39

Although this experiment will inevitably suggest itself, no such tests have yet been reported, to the knowledge of the author. Perhaps the problem was the difficulties in recognizing or adhering to the necessary experimental conditions.

3.5.1. Variation in τ_p with Diffusion Time /40

To study the variation in τ_p with a diffusion time t, gallium-diffused p^+n-diodes (10 mm diameter, 800 µm thick) were coated on both sides with gold layers 0.5 µm thick and placed in the diffusion oven at about 800°C for up to 180 hours. After the diffusion process, the diodes were quenched to room temperature as described above. The applied gold layers were removed in a $HCl-HNO_3$ solution. After 5-10 sec etching in CP6, the original current-voltage characteristic of the diodes was restored. This required the removal of a layer only a few µm thick from the silicon surface, and therefore produced no essential change in the diode dimensions. τ_p was measured before and after the diffusion process. The same diode was then again coated with gold, and τ_p was again measured after a longer diffusion time t. In this way, the decrease of τ_p with t could be determined.

During the first series of experiments, three diodes (1), (2), and (3) were diffused with gold at 785°C for different periods of time (max. 180 hours), and then the τ_p values were determined. Diodes (1) and (2) were from the same initial crystal D (Haldor-Topsoe 3-706-1), while diode (3) was prepared from another crystal E (Wacker 4K/ZN 325) with, however, similar properties. In every case, it was zone-pulled n-material with 30-50 Ωcm and 10^4 cm^{-2} mean dislocation density.

The experimental values are summarized in Table 4. For diffusion times longer than 100 hours, $\tau_p(t)$ converged to a minimum τ_{min} = 1.5 µsec. In Fig. 11, $\tau_{min}/\tau_p(t)$ is graphed vs. t on a log-log scale. Good agreement is found between (28) and the experimental curve. Nevertheless, the large scattering of the k_o values for the three diodes ($2.8 \cdot 10^{-2}$/hour, $2 \cdot 10^{-2}$/hour, and 10^{-1}/hour, respectively) is noteworthy. This scattering is unexpectedly large even for the values for diodes (1) and (2) from the same crystal.

3.5.2. Investigation of the Condition $\tau_p'' \approx \tau_p(t)$ /41

The experiments just described were also carried out on diodes with values of τ_p'' not much smaller than the corresponding $\tau_p(t)$ values after the first diffusion step. The diodes (6) from crystal B (Lopex 35442) and (7) from crystal F (Komatsu 36 4973) had τ_p'' values of 15.2 µsec and 16.0 µsec, respectively. After 1-hour diffusion at 775°C, these values diminished to 13.8 µsec and 8.9 µsec respectively.

The values of $\tau_p(t)$ determined in this way are listed as a function of t in Table 5. Figure 12 shows the values of τ_p'' calculated from (8) plotted against t. It can be seen that diode (7) reaches the minimum of 1.5 µsec after 180 hours, while diode (6) has not yet reached the saturation value after even t = 180 hours. The result on the Lopex crystal (lower value of k_o) is consistent with the finding in the activation analysis.

3.5.3. Influence of n^+- and p^+-Diffused Boundary Zones on $\tau_p(t)$

The previous considerations treated only problems associated with defined gold incorporation in simple, gallium-diffused p^+n-structures. The influence of additional n^+- and p^+-layers will now be investigated, in particular for p^+np^+- and $n^+p^+np^+$-thyristor structures. We prepared the corresponding n^+-layers by phosphorus diffusion, with a surface concentration of $C_o = 5 \cdot 10^{21}$ cm^{-3} and a diffusion depth $X_D = 10$ µm. For p^+-layers, the corresponding values were $C_o = 5 \cdot 10^{18}$ cm^{-3} and $X_D = 90$ µm.

Comparison of p^+n- and p^+np^+-Structures

Starting from ten p^+np^+-structures, five p^+n-diodes were prepared by lapping off the p^+-layers. The n-layer had a thick- /42
ness of $W_n = 500$ µm and a resistivity of $\rho_n \approx 50$ Ωcm; the repeated measurements of τ_p'' for these diodes furnished a mean value of 80 µsec. After deposition of gold from the vapor, these diodes were diffused together with the other five p^+np^+-structures at 850°C. Following the gold diffusion, the p^+np^+-structures were converted to p^+n-diodes, and $\tau_p(t)$ was determined for all ten elements. The same mean value of 1.9 ± 0.2 µsec was found for both element groups. This proves that additional gallium p^+-layers do not affect gold diffusion in any way.

The Influence of Gold Diffusion in Thin Diodes on One or Both Sides

Five further p^+n-diodes from the same diffusion series were now coated with gold on one side (the n-zone) and then diffused under the above experimental conditions. The resulting mean value for $\tau_p(t)$ was 2.0 ± 0.2 µsec. Comparing this value with the value of 1.9 ± 0.2 µsec for diodes diffused on both sides proves that $\tau_p(t)$ under the given experimental conditions does not depend on whether gold is diffused in on both sides or just one side. This experimental finding can be interpreted as follows:

In the center of the slice, gold concentration is independent of position (X). There must be just enough interstitial gold atoms present to guarantee the condition $C_I = C_I^+$ for all X at the end of the first diffusion phase. These interstitial atoms can be supplied from either one or both surfaces.

Comparison of $n^+p^+np^+$- and p^+n-Structures with Respect to Gold Diffusion

Finally, we will investigate the influence of an additional n^+-layer on the gold diffusion process. Theoretically, the /43
solubility and diffusion of a charged impurity atom in a semiconductor, in this case gold in silicon, are functions of the position of the Fermi level [40, 41]. This will be discussed in detail in Section 4.6 on the gettering of metallic impurities by highly doped n^+ surface layers. We will just mention here that gold has a high solubility but a low diffusion capacity in phosphorus-diffused n^+-zones. On the other hand, this is not observed for highly doped boron p^+-zones with surface concentrations of $C_o = 5 \cdot 10^{21}$ cm^{-3} [29]. This last fact agrees with the above result that gallium p^+-layers with a somewhat lower $C_o = 4 \cdot 10^{19}$ cm^{-3} do not affect the gold diffusion process. However, additional phosphorus or arsenic n^+-layers might be able to affect the gold diffusion process and bring about higher values of τ_p than those found for the p^+n-diodes investigated here. This should be checked.

For this purpose, 15 $n^+p^+np^+$-structures were divided into three groups. The first group was coated with gold on the n^+-layer (phosphorus), the second on the p^+-layer, and the third on both sides. They were then diffused under the conditions mentioned above. Subsequently, lower values of $\tau_p(t)$ were found only when the p^+-layer was coated with gold. In the other case, the final values of τ_p were unchanged.

In summary, it can be stated that gold diffusion in thin p^+n-silicon diodes yields values of $\tau_p(t)$ after diffusion which are independent of whether:

a) gold is diffused into the slice from one or both sides,

b) an additional gallium p^+-layer is produced resulting in a p^+np^+-structure,

c) there is also a phosphorus or arsenic n^+-layer, so that an $n^+p^+np^+$-structure is involved.

Higher values of $\tau_p(t)$ are observed only when existing /44
interstitial gold atoms can only reach the center of the n-zone by diffusing through the n^+-layer. This is a consequence of the high solubility and low diffusion capacity of gold in the n^+-layers.

3.6. Application of the Results to Control of Minority Carrier Lifetime

The irreproducibility of lifetime control via gold diffusion was attributed here to the nonuniformity of the values of k_0 and V_Q for the different crystals. As already shown in the previous sections, this manifests itself in the scattering of the $\tau_p(t)$ values for finite diffusion times (cf. (24) and (28)). Unfortunately, the present technological state of single silicon crystal growing does not permit the preparation of silicon with a higher degree of homogeneity relative to k_0. Hence, it is still not possible to specify diffusion conditions with respect to time and temperature to obtain a specific value of $\tau_p(t)$. One possible solution is the selection of diffusion times long enough so that $\tau_p(t) \approx \tau_{min}$. This can be correlated via the diffusion temperature. However, the experiments show that diffusion times of more than 100 hours at temperatures of 800°C are required, and a simple diffusion process all the way to saturation is therefore time-consuming. The effective diffusion time can nevertheless be reduced to a reasonable duration when the following two-stage diffusion process is employed, using silicon of <u>low</u> crystalline perfection.

Let the desired value of $\tau_p(t)$ be given by $\tau_{min}(T_1)$. The diodes are then diffused for a short time (e.g. 1 hour) at the temperature T_2, where $T_2 > T_1$. This is followed by a longer diffusion process at T_1, in order to reach $\tau_{min}(T_1)$ with the necessary precision [30]. The feasibility of this two-stage process is shown in the figures given in Table 6. Six diodes from five different crystals with widely varying τ_p" values were first gold-diffused for 1 hour at 820°C, and then for another 90 hours at 780°C. Except for diode 4 (Lopex, low k_0), the same minimum value of 1.5 ± 0.2 µsec was achieved throughout. /45

4. High Carrier Lifetime in Gallium-Diffused Silicon Structures /46

4.1. Introduction

In this section we will discuss the technological problems associated with the control of high hole lifetimes in the n-type zones of gallium-diffused p^+n-structures. By "control" we mean either

a) maintaining an originally high value of τ_p, or

b) recovering this high value of τ_p

by means of gettering processes. The necessity of achieving high τ_p values in certain silicon power components such as high-voltage thyristors has already been discussed in Section 1. As a special

example, the practical requirements on the forward characteristic of 4-kilovolt thyristors for HVDCT applications can be employed. Such a component requires an n-base thickness $W_n \approx 500$ µm, in comparison to $W_n \approx 100$ µm for a 1200-V element. Let the ratio $\sqrt{(D_p \tau_p)}/W_n$ be a quality factor for the forward characteristic. Then τ_p for the 4-kV element must be 25 times as large as that for the 1200-V element, for which $\tau_p = 4$ µsec is a representative value. Hence, values of τ_p on the order of 100 µsec are required for the 4-kV element. Similarly high carrier lifetimes are often needed in other types of silicon components prepared with conventional boron and phosphorus diffusion technologies, such as low-frequency power transistors and p-i-n diodes for utilization as nuclear-radiation detectors.

In view of the importance of this problem, the present work on experiments in recovering high carrier lifetimes by means of gettering processes for the preparation of silicon components should be of interest for communication engineering.

The diffusion of gallium (instead of the more frequently employed boron) into silicon to produce p^+np^+-structures is a key process in the preparation of thyristors. In the technology of silicon components for communications engineering, boron is employed instead of gallium as p-doping principally because of its extremely low diffusion constants in SiO_2 layers. Hence, its application in the production of planar, passivated and integrated components. However, the masking capacity of an oxide layer relative to a p-doping element is not required in thyristor technology, and moreover, the required deep-diffused junctions ($\approx$100 µm) are more satisfactorily produced with gallium than with boron. The main reason for this is as follows: /47

Because of its relatively small ionic radius (in comparison to silicon), boron, in concentrations greater than about 10^{16} cm^{-3}, produces a contraction of the silicon lattice sufficient to generate linear lattice defects, the so-called misfit dislocations [31]. Consequently, a boron-diffused layer 100 µm deep shows a disturbed surface region 50-60 µm deep. This is an undesirable effect, since this p^+-layer is used for the formation of a base region in the $n^+p^+np^+$ thyristor structure.

On the other hand, gallium has an ionic radius comparable with that of silicon: the ratio of the Pauling covalent radius of gallium to that of silicon is 1.068, compared with 0.746 for boron [32]. Furthermore, the solubility of gallium in the silicon lattice at a fixed diffusion temperature is smaller -- e.g. at 1250°C, $4 \cdot 10^{19}$ cm^{-3} vs. $5 \cdot 10^{20}$ cm^{-3} for boron -- and this is another reason why fewer misfit dislocations are observed in gallium-diffused layers. Another advantage of gallium over boron is its somewhat larger diffusion coefficient.

Gallium is conveniently diffused into the silicon from the vapor phase in sealed-off quartz ampules filled with argon, with a Ga-Si alloy as the source. In spite of careful handling and high purity of the quartz parts and of the chemical etching and cleaning reagents, the value of τ_p generally drops by an order of magnitude during a 60-hour diffusion at 1250°C -- from several hundreds to several tens of microseconds. This drop in carrier lifetime is usually blamed on the indiffusion of rapidly diffusing heavy-metal impurities such as Au, Cu, and Fe. In this section (4), we will study the benefits and the capabilities of gettering processes for the elimination of such impurities, particularly the gettering action of phosphorus-diffused surface layers, and possible gettering mechanisms. A new gallium diffusion technique is also presented, with which τ_p values of more than 200 µsec were achieved. Finally, results on the origin and the nature of the principal recombination centers penetrating the silicon during a conventional vapor-phase diffusion will be described. /48

4.2. Values of τ_p in Silicon after Gallium Diffusion from the Vapor Phase

4.2.1. The Gallium Diffusion Process

Before diffusion, the silicon was cleaned as follows: the slices obtained by sawing up single crystals were first lapped on both sides, in order to remove direct surface damage. This meant the removal of about 70 µm on both sides. Al_2O_3 powder (grain size 18 µm) in a petroleum oil of low viscosity was used as the abrasive. After the lapping, the slices were rinsed in trichloroethylene and washed in an ultrasonic bath, first in a detergent, in order to remove traces of the lapping oil, and then in deionized water ($\rho > 3$ MΩcm). Following this, the slices were rinsed in Pb-free HNO_3, in order to remove any macroscopic metallic surface impurities, and then in HF. After this cleaning process, about 100 µm of each side was chemically etched in CP6, and the slices finally rinsed in deionized water.

The diffusion itself was carried out in sealed-off high-purity quartz ampules. The source is prepared by alloying 5 mg metallic gallium on a silicon block. The gallium is placed in a small hollow in the silicon block. The alloying is carried out for 1 minute at 800°C in a H_2 atmosphere. The arrangement is depicted schematically in Fig. 13. The slices are stacked vertically with no space between them. Before the ampule was sealed, it was evacuated to 10^{-3} torr and filled with high-purity argon to a pressure of 1/5 atmosphere, resulting in a pressure of about 1 atmosphere at the diffusion temperature of 1250°C. The gallium vapor pressure at this temperature is 10^{-2} torr. This is sufficient to produce a surface gallium concentration /49

in the silicon of about $5 \cdot 10^{18}$ cm^{-3}. After a 60-hour diffusion, the silicon slices were cooled to 400°C at a rate of 1°C/min, and then quenched to room temperature. The depth of the resulting gallium-diffused layer was on the order of 90 µm.

This diffusion process makes it possible to economically produce deep-diffused p^+np^+-structures, which consistently possess an ideal reverse voltage characteristic after the necessary polishing of the edge of the slice. In addition, the slices retain their polished surfaces so that further diffusion steps can be carried out without the necessity of another surface treatment.

4.2.2. Values of τ_p after Diffusion

As already indicated, the values of τ_p after gallium-diffusion $\tau_p(Ga)$ measured by the diode retention time method are considerably shorter than the carrier lifetimes measured in the original specimens. In order to discover the material-related factors which could influence $\tau_p(Ga)$, 16 experiments were conducted with diffusion for 60 hours at 1250°C in each case. The results of this exhaustive study can be summarized as follows:

1. The mean values of $\tau_p(Ga)$ were between 5 and 60 µsec. /50

2. Apart from a few exceptions, $\tau_p(Ga)$ is roughly equal for slices of identical dimensions but from different silicon crystals, when they are diffused at the same time. By "different silicon crystals," we mean zone-pulled crystals with different dislocation densities ($N \approx 0$ to 10^4 cm^{-2}), different resistivities ($\rho \approx 10$ to 200 Ωcm), and crystals with the same specifications but from different manufacturers.

3. The values of τ_p within a crystal can vary by up to ±20% from the mean value. There was no well defined correlation of this scattering with the resistivity or with the dislocation density.

4. Diffusions which were repeated after a certain period (months) with the same material and under identical diffusion conditions yielded different values of $\tau_p(Ga)$.

From these results, we can draw the following conclusions:

1. The uniformity in the values of $\tau_p(Ga)$ for different but simultaneously diffused crystals suggests that the recombination center -- or the centers responsible for the decline in τ_p -- penetrated the silicon during the diffusion process from

the outside. On the other hand, this is incompatible with the results that a substantial number of centers are formed during the diffusion process or the high-temperature process due to internal rearrangement processes. For completeness, however, it should be pointed out that such internal center formation has been observed in silicon. Oxygen, which, as an impurity in crucible-pulled crystals, is usually electrically inactive, forms a donor complex SiO_4^+ when the silicon is tempered at 450°C. This donor then supplies a recombination center with an energy level 0.16 eV below the conduction band [33].

2. τ_p(Ga) is not appreciably affected by the presence of incorporated dislocations. /51

3. However, the scattering in τ_p(Ga) shows that the distribution of indiffused centers can be affected by the presence of other lattice defects such as dislocations and vacancy clusters.

4. Obviously, there is no connection between τ_p(Ga) and the original value of the carrier lifetime in the silicon crystal.

4.3. Phosphorus Gettering of Ga-Diffused Structures

In this section, we will discuss the results of experiments in increasing τ_p(Ga) by an additional phosphorus gettering process. Götzberger and Shockley [34] have already reported on the application of surface layers of phosphorus silicate glass and other glasses to improving the current-voltage characteristic of silicon p-n diodes. They observed that poor diode characteristics, such as high leakage currents and "soft" characteristic curves, can be considerably improved by growing phosphorus silicate layers on the silicon at elevated temperatures (about 1000°C). They assumed that the poor characteristics were associated with the formation of segregations of metallic impurities in the space charge regions. These impurities might have diffused into the structures during manufacture. It was proposed that phosphorus silicate layers be employed as getters for these metals, the getter action being produced by the higher solubility in the silicate layers, which are liquid at the gettering temperature. It is known that metallic impurities in silicon have small distribution coefficients $k(\sim 10^{-5})$, where k is defined as the ratio of the thermal-equilibrium concentrations of dissolved impurities in the solid and liquid phases of a material.

Therefore, it was assumed that the metallic impurities can be removed or gettered from the silicon volume, in that they would diffuse to these surface layers and dissolve in them. /52

The possible gettering action of phosphorus silicate layers was studied in this work for two reasons. First, one would expect that removing metallic impurities from Ga-diffused p^+n diodes would make it possible to obtain higher values of τ_p. Second, the phosphorus diffusion process required to produce such layers can be used in every case after the generation of the p^+np^+ structure to form the n^+ layer in the next step in the manufacture of $n^+p^+np^+$ thyristor structures. It is therefore possible to carry out an optimized gettering process and the formation of a diffused n^+ layer in a single step.

4.3.1. The Phosphorus Gettering Process

The phosphorus gettering or diffusion process was carried out in a conventional open diffusion system in a three-zone diffusion furnace. The slices were -- as shown in Fig. 14 -- stacked vertically in a quartz boat. A small quartz vessel containing about 0.5 g dry P_2O_5 was placed in a preheating furnace and held at a temperature of 250°C to 300°C. Under these conditions, the P_2O_5 sublimes and is carried along in the stream of dry argon, which flows at a rate of 60 ℓ/hour over the silicon slices, which are heated to the gettering temperature (800° to 1250°C). The P_2O_5 reacts with the natural oxide layer on the silicon surface, forming a phosphorus silicate. Elementary phosphorus diffuses out of this layer into the silicon, resulting in a highly doped n^+ layer.

The phosphorus silicate layer has e.g. a thickness of 0.6 µm after 1 hour at 1250°C, while the corresponding n^+ layer is 11 µm thick. Once this process was completed, the slices were slowly cooled to 600°C at a rate of 1°C/min, in order to avoid thermal stresses, and then removed from the furnace.

4.3.2. Values of τ_p after Phosphorus Gettering

The influence of phosphorus gettering on τ_p was studied for gallium-diffused diodes prepared in different diffusion series and at different points in the silicon. The gettering temperatures were varied between 800 and 1250°C and the times between 1/2 hour and 10 hours. The results of these studies are summarized in the following:

1. The change in the values of τ_p after gettering -- expressed by the ratio $\tau_p(get)/\tau_p(Ga)$ -- was found to be independent of gettering conditions (temperature and time) above 900°C. $\tau_p(get)/\tau_p(Ga)$ fluctuates between the values unity, i.e. no getter influence, and 4. Frequently, the values were between 1 and 2. No decrease in τ_p in gettering was observed. On the other hand, it was not possible to recover τ_p values even

approaching the original order of magnitude. This would not be achieved until $\tau_p(get)/\tau_p(Ga) \approx 10$.

2. $\tau_p(get)/\tau(Ga)$ is independent of the zone-pulled silicon starting material employed, in particular, of its resistivity and dislocation density.

3. In the $\tau_p(Ga)$ range studied, between 5 and 40 µsec, $\tau_p(get)/\tau_p(Ga)$ was independent of $\tau_p(Ga)$.

4. Repeated gettering did not result in any appreciable change in $\tau_p(get)$.

Hence, this phosphorus gettering carried out with the intention of achieving higher values of τ_p in gallium-diffused components was only slightly successful. In order to understand better the results obtained in this case, further studies were conducted. In particular, the following questions had to be investigated:

a) Does gettering of metallic impurities through phosphorus diffusion actually take place?

b) If this is the case, what are the gettering mechanisms involved?

4.4. Phosphorus Gettering of Gold and Copper

Gold and copper are typical representatives of the rapidly diffusing heavy-metal impurities in silicon, which can act as effective recombination centers. Because of their large capture cross sections for thermal neutrons and the favorable half-lives of their radioisotopes ($\approx$2 days), they can also be conveniently detected by means of neutron activation analyses. It might also be true that the results of a study of possible gettering mechanisms of these metals through phosphorus layers are directly applicable to other heavy metals as well, such as Fe, Ni, Mn, W, etc.

4.4.1. Experimental Methods [35]

Silicon slices 900 µm thick, obtained from a crystal with a resistivity of 50 Ωcm and a dislocation density of 10^4 cm^{-2}, were tempered in a quartz ampule for 60 hours at 1250°C. From /55 prior experiments, it was known that under these conditions, gold and copper diffuse into silicon in sufficient amounts to be measured by activation analyses. Under the chosen experimental conditions, the gold concentrations increased from 10^{10}-10^{11} cm^{-3} (silicon starting material) to 10^{12}-10^{13} cm^{-3}, and

the copper concentrations from 10^{11}-10^{12} cm^{-3} to 10^{13}-10^{15} cm^{-3}.

The slices were then gettered for 1 hour at temperatures between 800°C and 1200°C by the method described in Section 4.3.1. The gold and copper concentrations in the phosphorus silicate and in the underlying phosphorus n^+-layers and also the concentrations of the heavy metals remaining in the slices were then determined by means of neutron activation analysis. This was done in the following manner:

After gettering, the slices were irradiated for 3 days in a neutron flux density of 10^{13} $cm^{-2}sec^{-1}$, in order to activate the impurities. Thereupon, the concentrations of the metals in the phosphorus silicate layers were first determined after these layers had been dissolved in HF. Next, the underlying n^+-layer was etched off in steps of 1 µm with a buffered CP6 etchant until there was no longer any detectable radioactivity in these layers. The remaining segment of the slice was dissolved in CP6. The concentrations in the various layers were determined from the total amount of analyzed metal and the known layer thickness. The layer thicknesses could be calculated by weighing the slice before and after etching.

4.4.2. Gold and Copper Concentrations in Phosphorus Silicate Layers

According to Götzberger and Shockley [34], the getter action of the phosphorus layers is largely determined by the high solubility of the metals in these layers, which in [34] /56
were assumed to be liquid at the gettering temperature. The subsequently published phase diagram of the phosphorus silicate system confirmed the liquid state of this system above about 1000°C [36]. The gold and copper concentrations found in phosphorus silicate layers are listed in Table 7 as functions of gettering temperature. The concentrations found are practically independent of temperature. It is about 10^{16} cm^{-3} for Au and about 10^{18} cm^{-3} for Cu. There is no sign that these metals are more soluble at temperatures above 1000°C, when the layers are liquid. The liquid state of the system is thus not a prerequisite for the gettering action. As will be investigated later, it is more likely that the gettering action is due to the appearance of chemical compounds between the metals and the phosphorus.

4.4.3. Concentrations in the Slices after Gettering; Getter Relaxation Time

The gold and copper concentrations measured in the slices after gettering are likewise listed in Table 7. It was found

that the gold concentration decreases with increasing temperature up to the detection limit of $2 \cdot 10^{10}$ cm^{-3} at 1200°C. On the other hand, the gold concentration after the gettering at 800°C was always still comparable with that in non-gettered control specimens. In all cases, the copper concentration was below the detection limit of about 10^{12} cm^{-3}, i.e. at least an order of magnitude lower than that in the control specimens. Hence, effective gettering of these metals had taken place.

At this point, it is useful to estimate the order of magnitude of the times required to remove a specific impurity by a surface gettering process.

Let the concentration of the impurity at the surface be C_o, the thickness of the silicon be W. We assume that getter layers are applied to both surfaces of the slice in such a fashion that the boundary conditions expressed by (30a) are satisfied. The outward diffusion, i.e. the gettering, is given by (30) and characterized by the diffusion constant D. /57

Then the diffusion equation

$$\frac{\partial C(x,t)}{\partial t} = D \frac{\partial^2 C(x,t)}{\partial x^2} \tag{30}$$

with the boundary conditions

$$C(0,t) = C(W,t) = 0 \text{ for } t > 0 \tag{30a}$$

and the initial condition

$$C(x,0) = C_o$$

has the solution [37]

$$\frac{C(x,t)}{C_o} = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{1}{2n+1} \exp\left\{-\left[\frac{(2n+1)\pi}{W}\right]^2 Dt\right\} \sin\frac{(2n+1)\pi x}{W}.$$

The average concentration $\bar{C}(t)$ is:

$$\bar{C}(t) = \frac{1}{W}\int_0^W C(x,t)\,dx = \frac{8}{\pi^2} C_o \sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} \exp\left\{-\left[\frac{(2n+1)\pi}{W^2}\right] Dt\right\} \tag{31}$$

For $\overline{C}(t)/C_o \leq 0.8$, (31) has a simple approximate solution of the form

$$\frac{\overline{C}(t)}{C_o} \approx \frac{8}{\pi^2} \cdot e^{-t/t_o} \qquad (32)$$

where $t_o = W^2/\pi^2 D$ has the significance of a getter relaxation time. When the getter time $t = t_o$ has passed, (32) states that about 2/3 of the initial amount of impurities has been gettered. For this purpose, we consider the gettering of a 500 µm thick silicon slice at 1000°C. The effective diffusion coefficients D at this temperature are, according to [38], $D(Cu) \approx 5 \cdot 10^{-5}$, and $D(Au) \approx 5 \cdot 10^{-7}$ $cm^{-2}sec^{-1}$. With W = 500 µm, we obtain $t_o(Cu) = 5$ sec and $t_o(Au) = 500$ sec. In accordance with the model employed, the getter relaxation times are therefore comparatively small and, in particular, only on the order of minutes for the cases of practical interest. This is consistent with experience. However, it should be noted that the solution of the diffusion equation (30) for heavy metals in silicon does not hold strictly, since the diffusion process is more complicated than expressed in (30), as discussed in detail in the preceding Section 3. This will also be clear from the diffusion profiles of the metals in the gettered slices (see next section). /58

4.4.4. Concentrations in the n^+-Layers

The phosphorus-doped n^+-layers were etched off in steps of 1 µm -- as described above -- the gold and copper profiles obtained after 1 hour and temperatures of 1200 and 1100°C are depicted in Fig. 15. Figure 15 also shows the associated phosphorus diffusion profiles obtained from control specimens with the aid of stepped removal and layer resistance measurements. Figure 15 shows that the solubility of these metals is very large -- particularly in the n^+ regions where the phosphorus concentrations are greater than 10^{20} cm^{-3}.

Furthermore, in comparison to the corresponding silicate layers, higher absolute gold and copper amounts were found in the n^+-layers, which increase steadily in thickness at high temperatures.

The ratio V of the quantities of metal in the n^+-layer and in the silicate layer is plotted in Fig. 16 for various gettering temperatures. These results appear to indicate that, as their thickness increases, the n^+-layers become more effective getter "sinks" than their corresponding silicate layers. /59

4.5. Preliminary Conclusions

From the results presented in Section 4.4.3, it is clear that the phosphorus gettering process, even when conducted at relatively low temperatures below 1000°C, is very effective for fractionating out traces of gold and copper, which have entered the silicon during a diffusion treatment. (A similar result for the phosphorus gettering of iron in silicon has been reported by the author and coworkers in another place [39].) Another interesting finding, which supports this conclusion, was obtained by gettering silicon diodes doped with gold and diffused with gallium. Gold was diffused into ten p^+n diodes at 830°C for 1 hour. This lowered τ_p from 5.4 ± 0.3 µsec to 2.6 ± 0.2 µsec. The diodes were then gettered with phosphorus for 1 hour at 1100°C. This resulted in an increased value τ_p = = 5.7 ± 0.4 µsec. For ten control diodes of the same crystal and the same diffusion series, on the other hand, the same value τ_p = 5.7 ± 0.4 µsec was achieved by gettering alone. This means that the indiffused gold was completely gettered. This clearly suggests that the phosphorus gettering process is effective not only for Au, Cu, and Fe, but also for other heavy-metal impurities such as Ni, Mn, Co, W, etc.

The results obtained by phosphorus-gettering of gallium-diffused structures and described in Section 4.3.2. by the ratio $\tau_p(\text{get})/\tau_p(\text{Ga})$ now appear comprehensible if it is assumed that /60
during the gallium diffusion treatment, two types of recombination center penetrate into the silicon, namely a heavy metal, which, however, can be gettered by phosphorus layers, and perhaps a slowly diffusing nonmetallic impurity, which cannot be gettered. The relative concentrations of these impurity types probably vary statistically from one diffusion series to the next, but the concentration of the impurity which cannot be gettered is likely to be higher. References to the possible origin and nature of these latter impurities will be given later.

4.6. Gettering Mechanisms for Gold in Phosphorus-Doped Layers

In this section, we will attempt to discover the gettering mechanism for gold in silicon effective in the presence of highly phosphorus-doped layers. We conducted the following experiment.

Gold was diffused into eight silicon slices for 20 hours at 900°C. Then the slices were quenched to room temperature. After reheating at 1160°C diffusion was carried out in a phosphorus, arsenic, or boron atmosphere, in each case long enough to produce roughly equally thick diffusion layers. After removal of the diffused surface layers, the gold concentration in each of two slices was determined by means of neutron

activation analysis. The results are compiled in Table 8. Table 8 also contains the results of studies on the influence of a 6-μm thick gallium surface layer on the gold concentration. The table shows that phosphorus- and gallium-doped layers exert a considerable gettering effect, while arsenic-doped layers show only a slight effect, even for longer gettering times, and boron-doped layers show no gettering action at all. Furthermore, it should be mentioned that no increase in τ_p was observed when arsenic was used, in contrast to the result with phosphorus diffusion. /61

We will now discuss different gettering mechanisms which might play a role in the fractionation of metals -- particularly gold -- from silicon. These are (1) segregations at misfit dislocations, (2) solubility increases due to the electron-hole equilibrium, and (3) formation of ion pairs or compounds between atoms of the metal and the doping substance.

4.6.1. Segregations at Misfit Dislocations

As discussed in Section 4.1, the high surface concentrations produced in phosphorus or boron diffusion generate very "lattice-defective" surface layers containing a high density of misfit dislocations. It is known that the elastic strain fields of dislocations act on a number of metallic impurities. This can lead to the formation of segregations. The origin of segregations at misfit dislocations could therefore be viewed as a possible gettering mechanism in silicon. However, the absence of any gettering of gold in boron-doped silicon slices (see Table 8) precludes any such mechanism in this particularly important case.

4.6.2. Solubility Increase Due to Electron-Hole Equilibrium

Both Reiss et al. [40] and Shockley and Moll [41] have discussed the question of how the solubility of a charged impurity in a semiconductor depends on the position of its different levels relative to the Fermi level. In thermodynamic equilibrium, the concentration of neutral impurities is independent of the electrical potential. Charged metallic impurities, on the other hand, can have higher solubilities in highly doped surface layers than in weakly doped volumes. If the impurities are also sufficiently mobile, a gettering action can result. /61

We will now discuss this possibility for gold in silicon in greater detail.

Gold at lattice sites has three charge states (Au^-), (Au^x), (Au^+), as depicted schematically in Fig. 17. The probability w(E, N) that an impurity will have the energy E and N extra

electrons is given [42] by

$$w(E, N) = \text{const.} \exp\left(\frac{E_F N - E}{kT}\right)$$

For the three states of the gold atom, we thus obtain, neglecting level degeneracies

$$\begin{aligned} w^X &= \text{const.} \exp\left(\frac{E_F - E_D}{kT}\right); \quad w^+ = \text{const} \\ w^- &= \text{const.} \exp\left(\frac{2E_F - E_A - E_D}{kT}\right) \end{aligned} \tag{33}$$

This implies

$$N^+ : N^X : N^- = \exp\left(\frac{E_D - E_F}{kT}\right) : 1 : \exp\left(\frac{E_F - E_A}{kT}\right) \tag{34}$$

where N^+, N^X, and N^- stand for the gold concentrations in the different charged states.

Assuming that N^X is independent of the Fermi level, we obtain for nondegeneracy

$$N^+ = N_i^+ \exp\left(\frac{E_{F_i} - E_F}{kT}\right) = N_i^+ \frac{p}{p_i}, \tag{35}$$

$$N^- = N_i^- \exp\left(\frac{E_F - E_{F_i}}{kT}\right) = N_i^- \frac{n}{n_i}, \tag{36}$$

/63

where the index i designates the corresponding quantities in the intrinsic material.

Taking (34) into account, the gold concentration $N = N^+ + N^X + N^-$ is found to be

$$N = N^X \left[\exp\left(\frac{E_D - E_F}{kT}\right) + 1 + \exp\left(-\frac{E_A - E_F}{kT}\right)\right] \tag{37}$$

It is immediately evident from (36) that the concentration of negatively charged gold should rise above the value in intrinsic silicon as soon as $n/n_i > 1$. This is the case in silicon at 1000°C for $n \gtrsim 10^{19}$ cm^{-3}. Direct experimental verification of the result found in (37) is not yet possible, however, since the gold levels E_D and E_A at high temperatures are not known. Perhaps it should also be remarked that a corresponding decrease of N^+ follows in n-doped zones. However, this effect is negligible since the initial concentration of gold in the positively charged state is likely to be small [29].

This mechanism must also be insignificant in gold gettering, since arsenic-doped layers -- in contrast to comparably phosphorus-doped layers -- show only a slight gettering action (see Table 8).

4.6.3. Formation of Ion Pairs and Compounds

The formation of ion pairs -- for example, between a negatively charged gold ion and a positively charged phosphorus ion -- can also raise the gold concentration in phosphorus-doped n^+ layers beyond the increase described in Section 4.6.2. Reiss et al. [40] first described the formation of such ion pairs, namely for Li^+ and B^- in germanium. Here is a possible pair reaction: /64

$$Au^X + e^- + P^+ \rightleftharpoons (AuP)^X , \qquad (38)$$

where e^- is an electron and $(AuP)^X$ a neutral pair. Using the symbol Cp for the pair concentration, the mass action law then implies

$$C_P = \text{const.} N^X n N_D^+ \approx \text{const } N^X N_D^2 \text{ for } n \approx N_D \approx N_D^+ \qquad (39)$$

Here n is the electron concentration and N_D the donor concentration. If such a mechanism is actually operative, the pair concentration should increase quadratically with the donor concentration. Relevant measurements have already been made:

Cagnina [43] investigated the solubility of gold in silicon highly doped with phosphorus or arsenic. His specimens had homogeneous doping concentrations up to $8 \cdot 10^{19}$ cm^{-3}. Gold was diffused into his specimens at 1000°C up to saturation. The results are depicted in Fig. 18. Although (36) demands that N^- be a linear function of n or N_D, the nonlinearity of N^- is immediately evident. In both cases, the solubility of gold is roughly proportional to the square of the donor concentration, as required by (39). In this connection, the somewhat higher solubility of gold in the phosphorus-doped specimens is also noteworthy. However, the difference is not large enough to explain the substantially more effective gold gettering of phosphorus-doped layers in comparison with arsenic-doped ones. Nevertheless, the effect appears plausible because of the different chemical affinities between gold and phosphorus on the one hand and gold and arsenic on the other.

Namely, Wolley and Stickler [44] found the compound Au_2P_3 in gold- and phosphorus-doped silicon, while stable compounds between gold and arsenic are unknown at higher temperatures [45].

Other metal-phosphorus compounds are known as well, including Cu_3P and FeP. We therefore assume that these metals as well are mainly gettered by compound formation. These considerations should also apply to the gettering action of silicate layers. /65

4.7. Gallium Diffusion in Silicon from Liquid Surface Layers

In this section, we will discuss a new technology for diffusion of gallium into silicon [46]. Liquid surface layers of metallic gallium on the silicon surface are used instead of the vapor phase as the diffusion source. At a diffusion temperature of e.g. 1250°C, the silicon-gallium alloy has a content of 60 wt-% silicon, as given by the phase diagram [45]. An initial gallium layer of thickness w_{Ga} will therefore dissolve at this temperature a silicon layer w_{Si}. According to the lever law valid for this phase diagram, we have:

$$40/60 = \left(\rho_D^{Ga} / \rho_D^{Si} \right) \cdot w_{Ga} / w_{Si} \qquad (40)$$

With $\rho_D^{Ga} = 5.95$ g/cm^3 and $\rho_D^{Si} = 2.33$ g/cm^3, we then obtain $w_{Si} \approx 4\, w_{Ga}$. During diffusion, there is a liquid layer of thickness $w_{Si} + w_{Ga}$ on the silicon, and one might therefore hope to exploit the possible gettering action of this layer and the simultaneous gallium doping of the underlying silicon to prepare p^+n-structures with high τ_p values. Investigations on this point will now be described.

4.7.1. Preparation of Gallium Layers

We prepared thin Ga layers in two different ways:

In the first procedure, we exploited the low melting point of gallium ($T_m = 29.8$°C). Above 30°C, gallium can then be mechanically distributed over a silicon slice warmed slightly to e.g. 60°C e.g. by means of a Teflon strip. In this way, a complete coating was obtained for layer thicknesses up to several µm, but the layers produced in this way do not have a uniform thickness. This can be improved by vapor deposition. Therefore, in most of the experiments, the layers were evaporated on. However, there were no theoretical differences between the two methods. /66

4.7.2 τ_p after Liquid-Phase Diffusion

We first check the values of τ_p obtained in p^+n structures with liquid-phase diffusion as compared with those delivered by conventional vapor-phase diffusions. For this purpose, four slices were cut from a zone-pulled silicon crystal 26 mm in diameter (50 Ωcm), lapped and chemically polished; the thickness was then 4 mm. On two of these slices, high-purity gallium (99.9999%, i.e. 6-9) was deposited from the vapor in a layer 6 µm thick. For the time being, no gallium was added to the other two slices. All four were inserted vertically with a 1 mm separation in an argon-filled quartz ampule, and diffused for 60 hours at 1250°C. The gallium on the surfaces of the coated slices then acted simultaneously as a vapor source for the as yet gallium-free slices. After diffusion, all slices were lapped on one side, in order to remove the p^+-layer. Three diodes were drilled ultrasonically from each slice, and τ_p values were then measured by the diode retention time method. Values between 220 and 350 µsec were obtained for the diodes diffused with liquid gallium, and values of 40 to 70 µsec for the diodes prepared by vapor-phase diffusion. These τ_p values proved to be reproducible at this magnitude as well. Hence, when liquid gallium was used, the recombination centers were either prevented from penetrating or else gettered by the liquid surface layer. /67

Recently, Muraoka, Kato, and Nakumura [47] also reported similarly high values of τ_p (100 to 300 µsec) in gallium-diffused structures (from the vapor phase). However, the practical importance of their process appears doubtful, since the requirements relative to purity of the materials and of the chemical reagents are very high.

4.7.3. Influence of Layer Thickness and Degree of Surface Coverage on τ_p

Next, we investigate more precisely the influence of layer thickness on τ_p. For this purpose, slices with vapor-deposited gallium layers 6, 3, and 1 µm thick on both sides, together with slices coated only on one side with 6 µm gallium, were diffused simultaneously in the same quartz ampule under the usual conditions. We obtained τ_p values of 100-300 µsec for the 6 µm layers, 60-200 µsec for the 3 µm layers, and 40-60 µsec for the 1 µm layers. The slices coated only on one side achieved values of 60 to 160 µsec.

These results, in particular the dependence of τ_p on layer thickness, support the hypothesis that the penetration of lifetime-reducing impurities is hindered, if not completely prevented,

by the liquid surface layers. An alternative explanation, namely, that existing but previously electrically inactive impurities in the silicon are gettered from the volume, is unlikely, because a gettering action for the small quantities of impurities should not depend to such a degree on the layer thickness. /68

4.7.4. Influence of Slice Preparation and Purity of Gallium on Values of τ_p

The values of τ_p after the liquid-phase diffusion proved to be insensitive to the purity of the gallium employed: the same high values of τ_p of 100 µsec were obtained even when a) the surfaces were rinsed in ordinary tap water, or b) the slice was exposed to the free atmosphere for several hours before the gallium layer was deposited. Neither did the utilization of 99.9% gallium, instead of 99.9999%, result in any reduction of the τ_p values. For this reason, liquid-phase diffusion as a step in production represents a considerable reduction in expense. A drawback of the method is the necessity of removing the gallium-silicon alloy layer before further diffusion steps are undertaken. In the previously investigated specimens, the wetting of the Si surface by the gallium layer also was often incomplete, so that the penetration depth of the alloy zone was subject to fluctuations. Nevertheless, we are convinced that the liquid-phase gallium diffusion in the preparation of p^+n-structures with very high τ_p values will result in noteworthy progress in silicon thyristor technology, when these difficulties have been overcome.

5. Measurements of the Hall Effect in Silicon for Determining the Concentration of Deep Impurities /69

5.1. Introduction

The detection methods for deep impurities, which are still effective in extreme dilution, are of great importance for a systematic manipulation of lifetime doping. The problem is to determine concentration (in the range 10^{10}-10^{13} cm^{-3}) and chemical association. In addition to the frequently employed -- but not successful for all elements -- activation analysis, mass spectrometry and electron spin resonance come into consideration for the impurity analysis. Apart from these relatively elaborate techniques, which generally cannot be employed in connection with production, photoconductivity [49, 50, 51] and capacitance-voltage measurements [52] and the Hall effect measurements [53, 54] discussed in this chapter can supply information on the impurities present.

Measurements of the Hall effect as a function of temperature furnish the relationship between charge carrier density and temperature. From this can be derived the concentration and perhaps even the energy level of the deep impurities. However, one is subject to the limitation that only the upper half of the forbidden band can be studied in n-type material, and only the lower half in p-type material. By comparing the measured activation energies with literature values, the element can be identified. The activation energies of the known impurities in silicon in the region of interest of the periodic table are listed in Table 9. This representation provides a better guide in the search for unknown impurities. In considering the Table, it is evident that the role of a large number of elements as impurities in silicon has not yet been investigated. The figures were taken from the survey work by Schibli and Milnes [48], and supplemented or modified by data from more recent research, e.g. for Pt [55], Co [56], Ag [57], S [58], and Na [59]. The majority of known deep impurities are found in the first and second subgroups, the sixth group, and among the transition elements.

To justify the decision to record Hall effect measurements as a function of temperature, we should remark that this procedure constitutes the first improvement over a simple resistance measurement, involving both the information to be obtained and the effort required. The procedure furnishes directly the charge carrier density and the type of conduction. Its sensitivity grows with decreasing impurity concentration, a particularly favorable circumstance in the investigation of low-doping-concentration silicon for power components. In combination with the resistivity, which can be measured with the same experimental instrumentation, the mobility of the charge carriers is obtained.

5.2. Construction and Operation of the Apparatus for Measuring the Hall Effect in Silicon Specimens at Temperatures of 100-480 K

5.2.1. Experimental Setup and Equipment

The measuring system employs the known d.c. steady field technique. By reversing specimen current and magnetic field, undesirable galvanomagnetic, thermomagnetic, and thermoelectric interference voltages are eliminated. The circuit is depicted in Fig. 19. In designing it, we started from the fact that relatively weakly doped silicon specimens were to be investigated, i.e. that the specimens could be extremely highly resistive particularly at low temperatures. To generate the magnetic field, we employed the precision magnet B-E15C8 from Bruker-Physik AG, Karlsruhe, with current-stabilized power unit,

precision class 10^{-5}, and externally controlled reverse switching unit. With a pole shoe diameter of 15 cm and an airgap width of 56 mm, it normally operates with a flux density of 0.845 T. The magnetic flux density is determined by means of an InAs Hall probe, the characteristic curve of which was previously recorded in calibrating magnets. /71

Resistivity and the Hall effect were measured by the method of van der Pauw [60]. The specimen current was drawn from a battery and its setting monitored with a digital ammeter (digital picoammeter 445, Keithley). In accordance with the resistivities of the Si specimens, the currents were between 10^{-4} and 10^{-9} A. The voltages were measured via a battery-driven impedence converter (Model 311 K, Analog Devices) with a digital voltmeter. The Si specimen in the cryostat and all of the remaining circuit were electrically shielded. The circuit was highly insulated, using teflon.

During the measurement, the specimen is situated in a teflon holder in a Dewar vessel under nitrogen. Its temperature is controlled by a gas stream drawn from a temperature-control unit (AEG). There is no electrical heating in the interior of the shielding. Temperature was measured with thermocouples on the forward and reverse sides of the specimen.

5.2.2. Preparation of the Specimens

The silicon specimens studied were 19-25 mm in diameter and 200-600 µm thick. The simplest preparation method is to weld ultrasonically four aluminum wires as contacts to the circumference of the etched Si wafers. The ohmic component in the Hall effect measurement can be kept small by precise adjustment of the contacts. The wire contacts satisfy very well the condition of "point shape." Since no special temperature treatment is required in the preparation, there is no danger that rapidly diffusing impurities will penetrate into the specimens during application of the contacts. The aluminum ultrasonic contacts worked well for relatively low-resistance n-type and p-type silicon ($\rho \leq 50$ Ωcm). The resistive behavior was reviewed in the temperature range under investigation. Other contacts, such as pressure contacts and wires cemented on with conducting pastes, proved to be unsuitable. /72

For highly resistive specimens, in particular for slices whose net doping consists of deep impurities, the utilization of alloy contacts with larger contact surfaces cannot be avoided. For n-type material, Au(Sb) contacts (alloy temperature 450°C] were used, for p-type material, aluminum-silicon eutectic

contacts (alloy temperature 750°C). No changes in the doping of n-type material could be observed even for very highly resistive Si slices (ρ = 11,000 Ωcm). As a precaution, the slices were briefly etched after the alloying, in order to remove any possible surface impurities. On the other hand, the higher alloying temperature for p-type silicon proved to be more critical. After the alloying, at least 20 µm had to be etched off particularly highly resistive slices, in order to remove more highly doped surface layers produced during the alloying. Experiments with Au(Ga) and Au(B) as contact substances did not yield satisfactory results on highly resistive p-type material. The alloy contacts were applied in the form of round wafers 3 mm in diameter. Foils 50 µm thick were employed. Aluminum wires were welded ultrasonically to the alloy contacts, the wires then serving to hook up the specimens to the measuring instrumentation via terminals. In order to prevent the "point shape" condition from being violated by the larger surface of these contacts, the specimens were cut to a cloverleaf shape with an ultrasonic drill in accordance with the proposal of van der Pauw [60].

5.3. Experimental Results

5.3.1. Determination of the Hall Coefficient

The Hall constant R_H is calculated from the experimentally measurable quantities Hall voltage U_H, specimen thickness d, specimen current I, and magnetic flux density B by means of the known relationship

$$R_H = \frac{U_H \cdot d}{I \cdot B} \tag{41}$$

In the case of electron conduction, the charge carrier density is obtained from R_H by means of /73

$$n = - \frac{r_n}{R_H \cdot e} \tag{42}$$

The so-called Hall coefficient r_n for electrons (the analogous statement holds for r_p, the Hall coefficient for holes) must be known in order to analyze the Hall measurements. With lesser demands on precision, or in analyzing exponential concentration changes, we can set r_n or r_p = 1, since the actual values are in this order of magnitude. For more precise studies, and in particular for our intention of determining a few deep impurities in addition to the primary doping of shallow impurities, a detailed treatment of the Hall coefficient cannot be avoided. It is after all the quantity with which the Hall mobility

$$\mu_H = R_H \sigma \tag{43}$$

can be converted to the drift or conductivity mobility

$$\mu_C = \sigma / n \cdot e \tag{44}$$

Namely

$$r = \mu_H / \mu_C \tag{45}$$

In Fig. 20, the Hall constant R_H measured for an n-type silicon specimen is plotted as a function of 1000/T. Aside from a not precisely known boron level (about $3 \cdot 10^{12}$ cm^{-3}), the specimen also contains phosphorus as a shallow impurity. Only the net doping -- about $4.1 \cdot 10^{13}$ cm^{-3} -- can be derived from the Hall measurements of Fig. 20. The Hall constant grows with increasing temperature by the factor 1.25 before the specimen becomes intrinsic. Since the range 1000/T = 10 to 1000/T = 2.6 is located in the exhaustion plateau of the shallow impurities, R_H is actually anticipated to be constant in this segment. From the fact that R_H increases with rising temperature, it can be inferred that this effect is not caused /74 by a change in electron concentration. Namely, this would imply that the concentration of electrons in the conduction band would drop with increasing temperature, quite in conflict with the thermodynamics. The temperature dependence found here for the Hall constant R_H at constant carrier concentration can be attributed to the temperature variation of the Hall coefficient r_n.

Theory [61] provides the following clues to the Hall coefficient: if, with great mobilities, a strong magnetic field is chosen in order to satisfy the condition $(\mu B)^2 >> 1$, the situation is simple, namely $r_n = 1$. For silicon, this condition cannot be achieved with reasonable experimental effort. In practice, the formulas for weak fields always hold. In that case, r_n depends in a complicated fashion on the band structure and the scattering mechanism of the semiconductor. With simplifying assumptions (spherical energy surfaces, scattering from acoustic lattice vibrations), theory yields $r_n = 3\pi/8 = 1.18$. Long [62, 63] and Long and Myers [64] made more precise calculations and comparisons with experiments for p-type and n-type silicon. However, for this theoretical calculation of the Hall coefficient, we must have more information than is a priori known for our specimens. One feasible solution, which was also employed by Long, consists of the following steps: the temperature variation of R_H is determined for a silicon specimen as clean as possible, e.g. from a zone-pulled starting crystal not yet subjected to any high-temperature process.

Assuming that the specimen contains no deep impurities (n = const.), R_H as a function of temperature (see Fig. 20) directly represents the temperature variation of r_n. From this measurement, we obtain only the relative change. The absolute value of r_n is found, in accordance with theoretical expectations, by the normalization $r_n(100\ K) = 1$. For a single specimen, one cannot be sure whether the assumption n = const. is really satisfied for $100 < T < 400\ K$ -- the Hall measurement will supply the first evidence on any existing deep impurities. Therefore, as insurance, a comparison of the temperature behavior of several specimens is required. The n-type Si specimen which exhibited the greatest increase in R_H between 100 K and the start of intrinsic conduction was then considered to be the cleanest specimen and used for the determination of r_n. Specimens from 20 different n-type starting crystals (net doping $1 \cdot 10^{13}$ to $5 \cdot 10^{14}$ cm^{-3}) were surveyed. In Fig. 21, the Hall coefficient r_n for n-type silicon calculated from these measurements is depicted as a function of 1000/T. The broken line is an extrapolated segment. For temperatures above 200 K, r_n obeys -- approximately --

/75

$$r_n = 1.20\ (T/300)^{0.22} \qquad (46)$$

The determination of the Hall coefficient r_p for hole conduction is analogous to the procedure for r_n. $R_H(T)$ was determined for clean p-type material with a boron doping in the range 10^{12}-10^{14} cm^{-3}. A typical result is reproduced in Fig. 22. Apart from the change of sign of R_H characteristic of hole conduction at the onset of intrinsic conduction, a decrease of R_H with increasing temperature is found even without the presence of deep impurities. In view of the relationship $R_H = r_p/pe$, r_p decreases as temperature rises. By comparing many specimens -- for p-type material, the specimen with the lowest relative change is the cleanest -- we obtain the correct temperature behavior of r_p. The result of these studies is depicted in Fig. 23. For the extrapolation into the region of mixed conduction (sign change of R_H), we obtain

$$r_p(T) = 0.785 - 7.03 \cdot 10^{-4}(T-320) \qquad (47)$$

A comparison of the $r_n(T)$ and $r_p(T)$ curves with the results of Long [62, 63] and Messier and Flores [65] shows a qualitative agreement, particularly in the temperature behavior. All investigations agree that r_n rises monotonically with increasing temperature, and r_p falls. The reference temperature of 300 K used in the literature cited appears inappropriate in view of the possible presence of deep impurities.

5.3.2. Studies on Phosphorus-Doped Starting Crystals /76

One focus of the studies was the typical starting material for power components. It can be characterized by the following information: 111-zone-pulled, n-conducting, doping: 10^{13}-10^{14} phosphorus atoms/cm^3, producer: Wacker-Chemitronic, Burghausen. Figure 24 depicts a segment from the concentration-temperature curve of five different n-type starting crystals. These curves are based on the $r_n(1000/T)$ curve shown in Fig. 21. The specimens contained a net doping between 2.6 and $3.4 \cdot 10^{13}$ cm^{-3} (corresponding to 150-200 Ωcm at room temperature). The exhaustion plateau of shallow impurities determines the transitions into intrinsic conduction, which are different for every crystal (the figures on the curves are crystal numbers). Here we obtain a qualitative characterization of the crystals in the statement: The sharper the bend, the fewer deep impurities there are in the crystal, i.e. the cleaner it is. On the right, toward lower temperatures, no further changes occur up to 100 K, corresponding to 1000/T = 10. Differences in the concentration-temperature curve are not observed until above 300 K. They can be attributed to the liberation of electrons from deep levels in the upper half of the forbidden band, and indeed to both donors and compensated acceptors. The concentration of these levels is in these crystals in the range 10^{12} cm^{-3}. This establishes that:

1. some of the recombination centers present in the finished component are already contained in the starting crystal;

2. even when they are produced in quantity at short time intervals by the same industrial process, the starting crystals contain deep impurities in varying concentrations.

Identifying the deep impurities via their activation energies is difficult, since in the n(1000/T) diagram of starting crystals, there are generally no definite stages, but instead gradual transitions. It must be assumed that several types of deep levels /77 exist side by side, and successively donate electrons as the temperature is raised.

To determine quantitatively the deep impurity concentration, which has as yet only been described qualitatively, we must examine more closely the temperature curve of the Hall constant R_H at the onset of intrinsic conduction. The Hall constant drops off in this region, due to the increase in electron and hole concentrations. However, the ionization of deep impurities is manifested in the decrease of R_H. In the

vicinity of the bend of the R_H (1/T) curve, the two effects superimpose. We must attempt to separate the component of electrons due to impurities. For the case of mixed conduction [53]:

$$R_H = \frac{\sigma_n^2 R_n + \sigma_p^2 R_p}{(\sigma_n + \sigma_p)^2} \tag{48}$$

σ_n and σ_p are the contributions of electrons and holes respectively to conductivity, R_n and R_p are the corresponding contributions to the Hall constant.

Using the Hall coefficients,

$$r_n = \mu_{Hn}/\mu_{cn} \tag{49}$$

$$r_p = \mu_{Hp}/\mu_{cp} \tag{50}$$

and the ratio of the Hall mobilities

$$C_H = \mu_{Hn}/\mu_{Hp} \tag{51}$$

we obtain

$$R_H = -\frac{r_n}{n \cdot e} \frac{\left(C_H^2 - \frac{p}{n} \cdot \frac{r_n}{r_p}\right)}{\left(C_H + \frac{p}{n} \cdot \frac{r_n}{r_p}\right)^2} \tag{52}$$

The concentration of majority carriers obeys

$$n = \frac{n_D^+ - n_A^-}{2} \cdot \left[\left(1 + 4\, n_i^2/(n_D^+ - n_A^-)^2\right)^{1/2} + 1\right] \tag{53}$$

and of the minority carriers

/78

$$p = \frac{n_D^+ - n_A^-}{2} \cdot \left[\left(1 + 4\, n_i^2/(n_D^+ - n_A^-)^2\right)^{1/2} - 1\right] \tag{54}$$

This will greatly complicate the direct conversion of the measured Hall constants R_H into the concentration of ionized impurities $n_D^+ - n_A^-$. We are confronted with an equation of fourth degree in n. The quantities r_n, r_p, C_H, and n_i appear in the coefficients. Because of the numerical uncertainties originating in the inexactness of the coefficients, and because of the great effort in computation, this method was not pursued any further. A quantitative comparison with theory was carried out in the reverse manner. Assuming that the concentration of ionized impurities measured in the exhaustion

plateau remains constant

$$n_D^+ - n_A^- = n_D - n_A = \text{const.} \qquad (55)$$

n and p are first calculated from (53) and (54), and finally R_H by means of (52), using experimental values for r_n, r_p, and C_H. In Fig. 25, using the example of an n-type starting slice, the experimental and calculated curves for R_H as a function of 1000/T is depicted in the region of the bend. The values in Fig. 21 and Fig. 23 were used for r_n and r_p, the relationship

$$C_H(T) = 4.37 \cdot (T/300)^{0.73} \qquad (56)$$

derived from (51), (59), and (61) for the mobility ratio, and the literature value [66]

$$n_i = 2.01 \cdot 10^{20} (T/300)^{1.5} \cdot \exp\left(\frac{0.605}{kT}\right) \qquad (57)$$

for the intrinsic conduction density.

The differences in the plateau are very small (they would hardly be expected to be otherwise because of the experimental value of r_n for a clean starting slice). In the bend, the measured R_H curve diverges downward due to the contribution of /79
electrons from deep impurities. In the region of intrinsic conduction, the measured and calculated curves are parallel.

For quantitative determination of the discrepancies from the theoretical curve, the ratio

$$v = R_H(\text{theoretical}) \; / \; R_H(\text{measured}) \qquad (58)$$

at 400 K corresponding to 1000/T = 2.5 was empirically selected. The comparison temperature of 400 K is certainly reasonable only in the doping range below 10^{14} cm^{-3}. Practically, we assume that all deep impurities have already surrendered their electrons at this temperature.

Crystals with v = 1 can thus be considered free of deep impurities in the upper half of the band, within the limits of measuring accuracy. Values of v greater than unity indicate impurities, while values of v less than unity should not occur.

Figure 26 depicts a histogram of v-measurements on 52 different n-type starting crystals (111-zone-pulled, diameter larger than 25 mm) with net doping between $2.7 \cdot 10^{13}$ and $1 \cdot 10^{14}$ cm^{-3}. They were crystals from industrial production in quantity (hence no test crystals) for diodes and thyristors with high

reverse voltage ratings. In the distribution of v-values, the steep rise at $y = 1$ is noteworthy as a fulfillment of theoretical expectations. For the majority of specimens, $1 \leq v \leq 1.15$. For a crystal with a doping of about $5 \cdot 10^{13}$ phosphorus atoms per cm^3, it is accordingly typical that it contains 10^{12} deep impurities per cm^3. This suggests the conclusion that these concentrations of deep impurities, which fluctuate from one crystal to the next, interfere with the control of carrier lifetime in the component production process. Nevertheless, the measurements furnished evidence that the initial level of deep impurities has fallen sharply in recent years, crystals with v-values greater than 1.15 being virtually nonexistent for more than 2 years. The cases given in the histogram (Fig. 26) belong to crystals pulled in the years 1965 to 1967 and provided by the manufacturer Wacker-Chemitronic for these studies. The reduction in the impurity level in the starting crystals is a result of technological improvements. Naturally, the question also arises of just how crucial the impurity level in the starting crystals actually is. This can only be discussed in connection with measurements on processed slices. For very large v-values, short carrier lifetimes will be anticipated after the component production process. According to these studies, there is no relationship between the v-values and the lifetimes measured on the crystals by the crystal manufacturer using the PCD technique (photoconductivity decay). /80

In addition to the temperature dependence of the Hall constant, the Hall mobility μ_{Hn} -- cf. (43) -- was determined for the starting crystals. In this case, it was found that the values for all crystals were closely bunched. At 300 K, a mobility of $\mu_{Hn} = 1520 \pm 20$ $cm^2V^{-1}sec^{-1}$ was obtained. At temperatures above 180 K, a decline was found, obeying the power law

$$\mu_{Hn}(T) = 1520\,(300/T)^{2.11} \qquad (59)$$

Aside from small differences in the vicinity of 100 K, which can be ascribed to different contributions of impurity scattering, no special features with respect to μ_{Hn} were observed. Therefore, a uniform quality has been achieved in zone-pulled n-type crystals with respect to mobility.

5.3.3. Studies on Crucible-Pulled Starting Crystals

Because of the high quality demands on material for power components, the studies have focused on zone-pulled crystals. In addition, measurements were taken on some crucible crystals

supplied by Wacker-Chemitronic. The following results were obtained: /81

The crucible material "as pulled" -- i.e. without any special thermal follow-up treatment -- exhibits the excitation of two levels in the concentration-temperature curve (see Fig. 27), Curve a). The plateaus in this example are at $n = 1.58 \cdot 10^{14}$ cm^{-3} (exhaustion plateau of phosphorus), $n = 1.84 \cdot 10^{14}$ cm^{-3} (exhaustion plateau of a donor with $E_{CD} = 0.12 \pm 0.03$ eV), and $n = 2.16 \cdot 10^{14}$ cm^{-3} (exhaustion plateau of an even deeper donor, E_{CD} cannot be determined more precisely). A rise can still be observed in this last plateau. Hence, there are still more impurities present. The almost identical differences between the concentrations of the plateaus suggests to us a double donor, probably involving the SiO_4 complex [48]. Determining the activation energies for small stages in the n(1000/T) diagram is connected with great uncertainties. The value $E_{CD} = 0.12$ eV was found by the technique given by Aliyeva and Tagirov [67] (cf. Fig. 30).

It is known that the SiO_4 complex reacts sensitively to temperature treatments. If the crucible crystal is subjected to a shock heat treatment, in which the temperature region between 500 and 300°C is traversed rapidly after repeated heating of the crystal, the levels largely disappear (see Fig. 27, Curve b). The electrically active fraction of oxygen has diminished. In the temperature region covered, there is no difference between the mobilities of crucible material and zone-pulled starting material, regardless of whether or not the latter exhibits dislocations.

5.3.4. Studies on Boron-Doped Starting Crystals

A series of p-conducting, boron-doped crystals were studied with the objective of obtaining experimental figures on the hole mobility μ_{Hp} and the Hall coefficient for holes as functions of temperature. Both quantities are required for the quantitative interpretation of the measurements on n-type material (cf. Ch. 5.3.2). /82 Investigation of the purity of the highly resistive p-type crystals later used as probes was likewise necessary (see Ch. 5.3.6).

An example for the Hall measurement on a relatively clean specimen has already been cited in Fig. 22. However, the Hall effect is less suitable for detecting deep impurities in boron-doped p-type material, since the sign change of R_H also takes place in the temperature region in which the deep impurities are ionized. The contribution of holes to the Hall effect is surpassed by the electron component. Therefore,

analysis is possible only when the formulas for mixed conduction are employed. In this case, taking the temperature variation of resistivity into account proves to be advantageous (see Fig. 28). In this diagram, resistivity is depicted as a function of temperature for three typical specimens. The broken curves were calculated theoretically with the following relationship:

$$\rho(T) = \frac{1}{e \cdot \mu_{Cp} \cdot \frac{n_A - n_D}{2} \cdot \left[(1+4n_i^2/(n_A-n_D)^2)^{1/2} (1+C) + 1 - C \right]} \qquad (60)$$

The depletion concentration of shallow impurities determined from Hall measurements was substituted for the net doping $n_A - n_D$. We resorted to figures of Ludwig and Watters [68] for the conductivity mobility μ_{Cp} and the mobility ratio $C = \mu_{Cn}/\mu_{Cp}$.

Overall, there is good agreement between measured and computed values. Curves a and b show clearly the depression of the resistivity maximum from the theoretical curve below 1000/T = 3, i.e. shortly before the onset of intrinsic conduction. These discrepancies in the ρ curve are caused by the excitation of deep impurities. No drop can be observed in Curve c.

The measurements show that there are deep impurities in detectable concentrations in p-type starting crystals as well. The attempt to correlate the lifetime values measured by the crystal producer (Wacker-Chemitronic) using the photodecay method with the concentrations of deep impurities led to the result, as with n-type starting material, that no relationship was discovered. /83

Determining the Hall mobility of holes μ_{Hp} yielded the power law

$$\mu_{Hp}(T) = 362\ (300/T)^{2.84} \qquad (61)$$

This result agrees well with literature values [61, 64], at least as far as the exponent. Rather large variations were observed from one crystal to the next in the absolute values of $\mu_{Hp}(300)$. μ_{Hp} tended to be larger for crystals with lower doping, e.g. $\mu_{Hp}(300) = 315\ cm^2V^{-1}sec^{-1}$ for $p = 2 \cdot 10^{14}\ cm^{-3}$, $\mu_{Hp}(300) = 356\ cm^2V^{-1}sec^{-1}$ for $p = 3.5 \cdot 10^{13}\ cm^{-3}$, and $\mu_{Hp}(300) = 430\ cm^2V^{-1}sec^{-1}$ for $p = 2 \cdot 10^{12}\ cm^{-3}$. These numbers can only reproduce a single observed tendency, and illustrate that uniform mobility values valid for all specimens can really not be assumed in an analysis of the ρ(T) curves. The differences between experiment and theory for Curve c of Fig. 28 likewise show this clearly. In the entire temperature range, the measurements show lower mobilities than assumed in the calculation.

5.3.5. Measurements on Si Slices after Different Processing Steps

In the manufacture of a component, it is unavoidable that further deep impurities will be introduced into the starting material. Rapidly diffusing impurities can penetrate throughout the entire volume of the crystal slice and in particular, can also change the carrier lifetime in the middle zone (base). We have studied slices of the same crystal both in the initial state and after high-temperature treatments such as Ga diffusion, oxidation, phosphorus diffusion, and gettering. A homogeneously doped specimen is always required for Hall effect measurements. /84
However, the base of the power component is generally thick enough (200-400 μm), to be prepared by lapping and etching and to be worked into a Hall specimen. The heavily doped zones are excluded from these studies. This is no problem, since the carrier lifetime in the base is the focus of interest.

Slices were studied both after individual steps and after passage through entire sequences of high-temperature steps. However, it is not technologically possible to measure a specimen and then subject it to a further step in the process, in order to measure it again. Thus the measurements always involve different slices, which do indeed belong to the same crystal, but which can exhibit differences in the doping of shallow and deep impurities, corresponding to the inhomogeneities of the crystal in the pulling direction.

The following results were obtained:

1. Ga diffusion (Si slices about 70 hours in Ga vapor in an argon atmosphere at about 1200°C) is always accompanied by penetration of relatively rapidly diffusing deep impurities. If the starting slices are generally found to have a v of at most 1.15, the slices after Ga diffusion have v-values in the base between 1.3 and 3 (see Fig. 29).

2. In the log n vs. (1000/T) diagram, there are steps on the upper plateau of which is superimposed the onset of intrinsic conduction (see Fig. 30). As the analysis shows, the levels which have penetrated are in our case donors of a single type.

3. The deep donors are homogeneously distributed in the base zones, but occur in different concentrations from one slice to the next. They act as recombination centers and affect the carrier lifetime.

Figure 30 depicts the results of Hall measurements on various slices of a crystal before and after Ga diffusion. The starting material (see Curve a) with v = 1.15 shows an

only slightly rounded-off transition from the depletion plateau to intrinsic conduction. The other three curves (b, c, d) were measured for base regions after Ga diffusion and show the concentration rises due to the ionization of the donors. The proof that it is actually donors which are involved at this level -- there could also be acceptors in the upper half of the band -- is given in Ch. 5.3.6. The concentration changes in Fig. 30 become large enough so that it is possible to calculate the activation energy of the deep donor with some precision. In the specimens, there are two donors with activation energies E_{CD1} and E_{CD2} in concentrations n_{D1} and n_{D2}. The first donor is the phosphorus of the initial doping. $E_{CD2} > E_{CD1}$, since the phosphorus donors are already completely ionized at the lowest temperatures of our measurement. Therefore $n = n_{D1}$ for the lower plateau. On the upper plateau, which is not so clearly defined experimentally because of the onset of intrinsic conduction, the second donor is exhausted at $n = n_{D1} + n_{D2}$. The concentration within the stage obeys the relationship [67]

/85

$$\frac{n(n - n_{D1})}{n_{D1} + n_{D_2} - n} = 0.5\, N_C \cdot \exp\left(\frac{-E_{CD2}}{kT}\right) \qquad (62)$$

derived from the neutrality condition and the mass action law (N_C = state density in the conduction band, $N_C \sim T^{3/2}$).

If $\log(T^{-3/2} \cdot (n^2 - n \cdot n_{D1})/(n_{D1} - n_{D2} - n)$ is plotted against 1000/T, the result is a straight line from the slope of which E_{CD2} can be calculated. This analysis is depicted in Fig. 31 using the example of Curve d from Fig. 30. With the plateau values $n_{D1} = 3.90 \cdot 10^{13}$ cm^{-3} and $n_{D1} + n_{D2} = 1.1 \cdot 10^{14}$ cm^{-3}, the activation energy is calculated to be 0.29 eV. With the same analysis, Curves b and c yield $E_{CD2} = 0.27$ eV. The question of the substance involved with this level will be discussed in Ch. 5.4.

The deep donor could be observed again and again in the course of the processing period of 3 years. It appeared regularly in gallium diffusion and was independent of the starting crystal employed. The concentration in the base regions was generally between $1 \cdot 10^{13}$ cm^{-3} and $7 \cdot 10^{13}$ cm-3. It varied from one slice to the next [rest of sentence cut off].

To attack the problem of whether this donor acts as a recombination center, lifetime and Hall effect must be successively measured for the same Si slice. The course taken at the beginning of the experiments, namely, to measure the Hall effect for some slices of a crystal and to prepare diodes for the lifetime measurement from other slices of the same crystal with the same treatment, no longer made sense because of the differences from one slice to the next.

/86

First, the decay time (see Ch. 6) was measured for the p-n-p structure present after Ga diffusion, and then the Hall effect was measured for the base region of the same slice. The results depicted in Fig. 30 show that the curves corresponding to the values of t_d measured in the slices are above one another. The cleanest slice (b) has the longest lifetime and the slice with the most deep centers (d) the shortest one. In Fig. 32 are collected results of decay time and Hall effect measurements on a series of four crystals (each represented by three slices), subjected to Ga diffusion together in a single ampule. We found that the reciprocal of the decay time t_d^{-1} rises markedly with increasing deep donor concentration (described here by $v = R_{Hcalc.}(400K)/R_{Hmeas.}(400K)$). There are rather large fluctuations, which can be explained by the presence of still other centers affecting lifetime, e.g. in the lower half of the band, which were not detected by our Hall measurements.

Gallium diffusion is the longest high-temperature process in the production of diodes and thyristors. It is the principal determining factor for the level of deep impurities in the component. According to our investigations, other steps such as oxidation in moist oxygen and the subsequent phosphorus diffusion do not cause any substantial changes in the impurity level. Hardly any, and in particular, no completely new deep impurities enter the silicon crystal. The Hall effect measurements on base regions of slices subjected to Ga diffusion and oxidation or Ga diffusion, oxidation, and phosphorus diffusion exhibited no measurable differences in comparison to slices into which only gallium was diffused. The v-values remained between 1.2 and 1.6, and the analyses of stages pointed to the deep donor already encountered. The fact that the fate of a specific slice cannot be followed through all the steps of the process and the fluctuations in deep-level concentrations from one slice to the next do not permit any more precise statements. /87

Not until gettering (see Ch. 4.3.1) do marked changes occur in the impurity level of the slices. The three concentration-temperature curves in Fig. 33 belong to the slices of one crystal. It shows the evolution of deep-impurity concentration from the initial slice through Ga diffusion and a concluding gettering. In this case, it can be observed that nowhere near all of the deep impurities incorporated are later removed. This is consistent with the finding that the initial lifetime can generally not be recovered by gettering. Measurements of the Hall effect can provide a check on the effectiveness of a gettering technique.

The changes in v-values in the process of producing a completely diffused thyristor, as followed here with the aid of the Hall effect (see Fig. 33), agrees qualitatively quite

well with the development of lifetime values. An attempt was made to discover some quantitative relationship. The lifetime and v-value had to be measured on the same slice. The lifetime measurement had to be made by means of a uniform procedure, applicable to all processing stages, in particular to starting slices. In the preparation of the specimens, in addition to the Hall specimen, some other slices 7 mm in diameter were drilled out ultrasonically from the starting slices or from the middle regions. From these slices, which had no p-n junctions, alloy diodes were then produced, for which the lifetime could be determined in accordance with Kingston [8]. The preparation of /88
the p-n junction by means of an alloy process offers the advantage that no long high-temperature treatment (e.g. Ga diffusion), with all the dangers of altering the specimen, is required. The result of this investigation is depicted in the diagram in Fig. 34. The reciprocal of lifetime is plotted against v, which gives us the impurity level. The starting crystals are in the shaded region near the origin. The values after Ga diffusion (triangles) are scattered over a wide region.

A tendency toward smaller v and longer lifetimes can be recognized for the gettered slices (circles). The fluctuations highlight the difficulties involved with quantitative studies on lifetime control. Nevertheless, the finding that specimens with high stages in the concentration-temperature diagram have short lifetimes can be considered confirmed.

5.3.6. Highly Resistive Si Slices as Probes for Penetrating Deep Impurities

There will not be favorable conditions for identifying by means of activation energies the impurities penetrating in high-temperature processes until the number of deep impurities is substantially larger than that of shallow ones. In a process control, it is not possible to generate at will unknown impurities in higher concentrations. However, there is another possibility. Highly resistive silicon slices with very slight shallow doping can be subjected to the process together with normal n-type Si slices. If unknown impurities enter, they can be very sensitively detected in the specimen slices. Probe slices from p-type and n-type silicon with net dopings of $2 \cdot 10^{12}$ cm^{-3} and less (resistivity at room temperature 6000-11,000 Ωcm) change their type of conduction upon penetration by even very slight concentrations. When p-type and n-type materials are used simultaneously, the result will show whether donors or acceptors are involved. For deep impurities with several levels, all activation energies /89
can be determined in a single experiment in favorable cases.

This method is a way of investigating the rapidly diffusing impurities penetrating from the outside, and determining their concentration. It does not detect deep centers brought in with the starting crystals. However, as long as the number of penetrating deep impurities exceeds to so great an extent the number present in the starting crystal (see Fig. 34), this method will be valuable. At the beginning of the research, radiochemical studies were also conducted, e.g. to identify the observed deep donor. However, although activation analysis is very sensitive for some elements, e.g. for Cu and Au, it fails for others, e.g. sulfur. The radiochemical studies proved useful in the sense that certain elements, namely the easily detectable ones, could be excluded when they did not occur. Nevertheless, it happened repeatedly that the Hall effect indicated deep impurities in appreciable concentrations, e.g. $5 \cdot 10^{13}$ cm^{-3}, while activation analysis did not give any sign of impurities.

The probe crystals were studied before utilization by means of the Hall effect and activation analysis. Except for boron, phosphorus, and some gold (close to the detection limit of $2 \cdot 10^{10}$ cm^{-3}), no impurities could be detected.

Probe slices (25 mm diameter, 540 µm thickness) of p-type silicon with a doping of about $2 \cdot 10^{12}$ boron atoms per cm^3 were subjected together with ordinary n-type starting slices to Ga diffusion. The probe slices naturally acquire heavily doped exterior p-type zones 50-100 µm thick, in accordance with the diffusion conditions. The following investigation of the middle zone showed that penetrating donors had reversed the type of conduction from p to n. The boron doping had been neutralized. Figure 35 depicts the concentration-temperature curve of such a slice. The net doping consists of deep donors with an activation /90
energy of $E_{CD} = 0.28$ eV. In the log n vs. 1000/T curve, we can see the three regions reserve with poisoning, normal reserve, and exhaustion plateau of the deep donor. The bend from the full slope to the halved value takes place at the level of the concentration of the neutralized boron. The number of acceptors in the specimen has not changed detectably. This result matches exactly the results found with n-type slices (see Ch. 5.3.5). Utilizing very highly resistive starting material brings out more clearly the influence of the deep donor.

In a total of ten diffusion tests, some with larger time intervals and with different probe crystals, similar results were always obtained. Only the concentration of the deep donor changed. The activation energy was repeatedly found to be 0.28 ± 0.01 eV. From the absolute determination of the mass action constants, the degeneracy of the donor level was calculated to be $g = 12$.

In further experiments, the source of the donor was sought and its behavior in various proces steps was observed. In a control test, highly resistive p-type and n-type slices were subjected to a high-temperature treatment together with ordinary n-type starting material without Ga source under otherwise identical conditions (cleaning, filling, and sealing of the quartz ampules, temperature program, and diffusion time). According to this test, the slices contained exterior Ga-doped p-type regions (detected by the Hall effect), but in much smaller concentrations than with a Ga source. In the interior, the donor was again found in concentrations of $7 \cdot 10^{12}$ cm^{-3}. This result shows that the quartz too is permeable to gallium at the high diffusion temperature of over 1200°C. A furnace used continually for Ga diffusion has a Ga level. It seems reasonable to seek the source of the deep donor outside the slice preparation and the Ga source. If the probe slices were provided with an oxide layer /91 before Ga diffusion, the penetration of the donor was prevented. By itself, the oxidation, like the gettering, caused no changes in the interior of the probe slices.

Figure 36 depicts the result of a series of tests in which all the steps for the production of a fully diffused thyristor were carried out on probe slices. Ga was first diffused into the slices. The result is, as shown by Curve a in Fig. 36, a donor concentration of about $1 \cdot 10^{13}$ cm^{-3}. After the subsequent oxidation of all slices, the oxide was completely removed from some (this corresponds to the windows in the oxide mask), while the oxide layer was left completely untouched in others. After the phosphorus diffusion, the concentration of the deep donor had risen further in the first group (see Curve b). However, no new type of impurity had penetrated. In the second group (see Curve c), on the other hand, the concentration had dropped sharply, in fact by a factor of 10. The combination of phosphorus and oxide layer thus shows an action similar to gettering (see Ch. 4.3.1).

5.4. Origin and Type of Prevailing Recombination Center after Vapor-Phase Gallium Diffusion

The results of these experiments show that the overwhelming majority of the recombination centers present in the finished power component were introduced into the silicon during gallium diffusion. The concentration of deep impurities acquired in this high-temperature treatment, which is the longest of the entire production process, can be only partially reduced in succeeding process steps. The centers penetrate into the slice from the outside. It is not, as one might surmise, thermal activation of latent substances or complexes in the original crystal. This is demonstrated by the masking effect of applied

oxide layers and the comparatively small influence of short high-temperature treatments.

In the past, the source of the impurities was primarily sought in the preparation of the slices and of the ampules. It was suggested that the centers might (a) be present on the silicon surfaces even before diffusion, or (b) be introduced into the ampule atmosphere via the argon, or (c) originate in the quartz of the ampules. It is true that these sources of impurities could be significant and certainly were in the past. However, even before the experiments discussed here, experience had shown that further improvements did not lead to any visible result.

Our results are evidence for the hypothesis that the significant source of impurities at the current state of technology lies outside the quartz ampule, perhaps in the ceramic protective tube or in the heating element of the furnace. The permeability of the quartz at high diffusion temperatures could play an important role.

The question of the chemical nature of the deep donor with $E_{CD} = 0.28$ eV, which surpasses in concentration all other deep impurities present, in particular those in the starting crystals, cannot be given a definitive answer. If it is assumed that the donor is not a complex, its position in the periodic table (see Table 9) can be localized. Besides the elements of the third and fifth groups, the alkali metals can be eliminated, because their levels are too shallow. The elements of group 1b (Cu, Ag, Au) can be ruled out on the basis of the negative results of the activation analysis. The strongest suspicion falls on the sixth group, particularly on sufur. This element satisfies three criteria: deep donor with $E_{CD} = 0.28$ eV, large diffusion coefficient ($D(1250°C) = 4.8 \cdot 10^{-8}$ $cm^2 sec^{-1}$ [69]), and not detectable by activation analysis. Besides the sixth group (S, Se, Te), suspicion falls on the second group (Mg, Ca, Sr), all elements for which little is known about diffusion properties and energy levels in silicon.

6. Determination of Lifetime in p^+np^+ and $n^+p^+np^+$ Slices by Decay Time Method /93

6.1. Introduction

In the preparation of thyristors and sometimes diodes as well, n-conducting silicon slices are converted into p^+np^+ or $n^+p^+np^+$ slices by means of diffusion processes. The lifetime -- of the charge carriers -- in these diffused slices -- particularly in the n-region -- plays a large role in determining the electrical properties of the component (e.g. recovery time and forward voltage).

If the lifetime after the diffusion processes is not in the desired range, it is raised or lowered by certain steps (gettering processes, gold diffusion) depending on the requirements. This leaves the conductivity doping essentially the same as it was after the diffusion processes.

The at least random-sample measurement of lifetime in p^+np^+ or $n^+p^+np^+$ slices -- before and after the steps of the procedure for exerting additional influence on lifetime -- is an important ingredient of production control. For example, the retention time method of Kingston [8] or the post-injection voltage method of Lederhandler and Giacoletto [7] can be employed for the lifetime measurement. Both methods require p^+n structures, and are therefore not nondestructive when applied to p^+np^+ or $n^+p^+np^+$ slices. The "sandwich photoconductivity" method of Bassett and Hogarth [70] makes it possible to take lifetime measurements on p^+np^+ structures. This method is not nondestructive either, because a p^+ zone must be partially removed in preparation for the measurement.

In this section (6), we describe a method for lifetime measurement on p^+np^+ slices, which we call the decay time method. This method is virtually nondestructive -- it is only necessary to etch the boundary of the slice. Moreover, this method is also applicable to $n^+p^+np^+$ slices, whenever the n^+ zone is imbedded in a p^+ zone.

6.2. Description of the Decay Time Method /94

6.2.1. Measuring Principle

The p^+np^+ slice, protected against light and held at a high temperature, is supplied with contacts in the form of pressed-on metal electrodes, and hooked up in series with a current measuring resistance R_M (Fig. 37). To this series circuit is applied a voltage $u(t)$, which is initially constant ($u = \hat{u}$) and then drops off exponentially with a time constant t_a. The current $i(t)$ flowing through the slice is observed. Under certain conditions, $i(t)$ passes through a minimum. The time interval between this minimum and the beginning of the voltage drop we call the "decay time t_d."

6.2.2. Temperature Dependence of Decay Time

With a given slice and fixed values of $\hat{u}$ and t_a, whether the current $i(t)$ passes through a minimum and the value of t_d (see Fig. 38) depend on the temperature θ. At low temperature ($\theta = 25°C$), there exists a minimum and therefore also a decay time.

However, the value of t_d is very small and depends in essence on the capacitances of the p-n junctions and on the measuring resistance R_M. With increasing θ, both t_d and the steady-state reverse current i_o rise: cf. the $i(t)$ curve for $\theta = 140°C$ with $t_d = 200$ µsec. At and above a "critical temperature" Θ, which in the present case is 148°C, $i(t)$ no longer passes through a minimum. That is, there is no t_d for $\theta > \Theta$.

6.2.3. Prior Results on the Connection Between Decay Time and Lifetime

Experimental and theoretical studies on the relationship between decay time and lifetime have already been reported [71]. We will briefly describe these results below.

6.2.3.1. Experimental Results

/95

A roughly linear relationship was found between decay time and lifetime for constants $\hat{u}$, t_a, and θ for slices of a specific type. By "slices of a specific type" we mean slices in which the exterior dimensions as well as the geometry and conductivity doping of the p^+np^+ or $n^+p^+np^+$ structure are virtually identical, but which can exhibit widely different lifetimes, and which can come from different crystals. The slices studied were gallium-diffused p^+np^+ slices, some of which were gettered after the gallium diffusion to increase the lifetime. The measuring temperature θ was chosen as high as possible, because there is a clear relationship between t_d and the lifetime only at sufficiently high θ. On the other hand, $\theta > \Theta$ must be satisfied. For gallium-diffused, and gallium-diffused and gettered p^+np^+ slices of a specific type, the critical temperature Θ lies within a spread of about 10° (at constant $\hat{u}$ and t_a). The measuring temperature θ was set about 10° under the lower limit of the Θ range. After the t_d measurement, diodes were prepared from the p^+np^+ slices (see Section 6.4), and the lifetime was measured by the retention time method [8]. Figure 43 shows results of recent studies of this type.

6.2.3.2. Theoretical Studies

The relationship between decay time and lifetime could, in principle, be interpreted theoretically. Nevertheless, the quantitative agreement between the theoretical and the experimental results was not and is still not satisfactory.

In Fig. 39, the conduction mechanism in a p^+np^+ structure is sketched, as far as required for comprehension of the following description. With the exception of lifetime (recombination and generation rates), we presume all parameters to be held constant (in particular, temperature, resistivity of the n-type region, $\hat{U}$, and t_a).

In the steady state (Fig. 39a), one p-n junction (J_1) is polarized in the forward direction and the other (J_2) in the reverse direction. Electrons produced in the space charge region of J_2 by thermal generation enter the n-type neutral region from the right. These electrons recombine completely in the n-type neutral region, as long as the emitter efficiency of J_1 is unity, which we assume for simplification. Holes flow across J_1 from the left into the n-type neutral region, where some of them recombine, while the remainder reach the space charge region of J_2. Taken as a whole, the n-type neutral region is an excess region, i.e. the charge carrier excess z is positive in the n-type neutral region. We define z by /96

$$z = \int \Delta n \, dx = \int \Delta p \, dx \qquad (63)$$

where the integral is to be taken from x = 0 to x = r_o or r(t). Δn and Δp are the electron and hole excess concentrations respectively, i.e. the difference between the electron and hole concentrations (n, p) and the thermal equilibrium values (n_o, p_o).

For the steady state, we obtain

$$z = \tau G w_o \qquad (64)$$

τ = lifetime in the n-type neutral region;
G = thermal generation rate in the space charge region of J_2;
w_o = width of the space charge region of J_2 in the steady state.

We assume that

$$G \sim 1/\tau \qquad (65)$$

This equation is based on the assumption that the recombination-generation processes in the n-type neutral region and the generation processes in the space charge region of J_2 take place at recombination centers of the same type, and that the concentration of recombination centers is homogeneous in the p^+np^+ structure. Equations (64) and (65) imply that the charge carrier excess z in the steady state is independent of the lifetime τ. /97

During the voltage drop (Fig. 39b), the current i(t) never goes through zero if the temperature is large enough (see Fig. 37). Therefore J_1 remains polarized in the forward direction and J_2 in the reverse direction. The J_2 space charge region shrinks and the n-type neutral region accordingly expands to the right. The width w(t) of the J_2 space charge region depends only on the voltage u(t). In the expansion of the n-type neutral region to the right, electrons along its right border (x = r(t)) are "heaped up" in a thin boundary layer. This neutralizes the space charge of donors, which shortly before still belonged to the J_2 space charge region. The number of electrons heaped up per units of time and area will be called the "neutralization rate" Q(t). Q(t) is given approximately by[1]

$$Q(t) = -n_D dw/dt \qquad (>0) \tag{66}$$

n_D = donor concentration in the n-type region;
w(t) = width of space charge region of J_2.

The electrons required for neutralization are supplied by virtue of the electron flux on the right side of the boundary layer between the n-type neutral region and the space charge region (x = r(t) + 0) being greater than on the left side of this boundary layer (x = r(t) - 0).

$$S_{r+0}(t) - S_{r-0}(t) = Q(t) \tag{67}$$

$S_{r\pm0}$ = electron flux[2] (particle current density) at the point x = r(t) + 0 and r(t) - 0, respectively. The electron flux is positive when the electrons flow toward the left. /98

Q(t) is independent of τ but $S_{r+0}(t)$ and $S_{r-0}(t)$ are not. If τ is small, the influx of electrons from the space charge region is sufficient to supply the ionization rate because of the high generation rate $G \sim 1/\tau$, i.e. $S_{r+0}(t) \geq Q(t)$. It is true that $S_{r+0}(t)$ decreases during the voltage drop, because w(t) becomes smaller, but the same also holds for Q(t), because dw/dt becomes smaller. If τ is large enough, the generation of charge carriers in the space charge region becomes so weak that $S_{r+0}(t) < Q(t)$ and therefore, by Eq. (67), $S_{r-0}(t) < 0$.

---

[1] Q and the capacitive current density j_C considered in [71] are related by the equation $j_C = qQ$ (q = elementary electrical charge).

[2] The electron flux S and the electron current density j_n considered in [71] are related by the equation $j_n = qS$.

That is, on the right side of the n-type neutral region, the electrons flow toward the right to the point r(t), because the n-type neutral region must, so to speak, make a contribution to the neutralization rate Q(t). This case is sketched in Fig. 39b.

The charge carrier excess defined by Eq. (63) is a function of time during the voltage drop. dz/dt depends on both z/τ (recombination excess rate in the n-type neutral region) and S_{r-0}. If τ is small, z(t) remains positive and falls monotonically to zero. If τ is sufficiently large, $S_{r-0}(t)$ is negative, as explained above. This means that the n-type neutral region itself must deliver some of the electrons required for its expansion toward the right. Since no electrons can flow in across the p-n junction J_1, the mean electron concentration in the n-type neutral region falls below its equilibrium value n_o. For neutrality reasons, the same holds for the mean hole concentration in the n-type neutral region, which falls below its equilibrium value p_o. Therefore, z(t) becomes negative. Then the generation of charge carriers, which causes z(t) to rise again, predominates in the n-type neutral region when the expansion on the n-type neutral region slows down in the further course of the voltage drop. The larger τ is, the lower the minimum of z(t), and the later it comes. However, z(t) cannot fall arbitrarily far. A lower limit for z(t) can be derived from the fact that in the n-type neutral region $n \geq n_D$, $p \geq 0$, i.e. $\Delta p \geq -p_o$ must always be satisfied. This leads to the condition

$$z(t) \geq -p_o r(t) \qquad (68)$$

r(t) = width of the n-type neutral region.

If z(t) were to reach the value given by (68), the neutrality condition in the n-type neutral region could only be satisfied by the space charge region of the p-n junction J_1 expanding, i.e. by J_1 changing over to the reverse direction. This would invalidate an essential prerequisite for the conduction mechanism described -- J_1 remains polarized in the forward direction. However, if the temperature is sufficiently high, this complication does not occur. One important reason is that p_o rises sharply with temperature. If the temperature is high enough, the lower limit $-p_o r(t)$ will then be so low that the conditions (68) remains satisfied. /99

Whether the current i(t) passes through a minimum, i.e. whether a t_d value exists, depends in essence on whether the charge carrier excess z(t) passes through a minimum. For small τ, no t_d value exists, because z(t) does not pass through a minimum. If τ is greater than a specific value τ^*, there exists the function $t_d = f(\tau)$, which first rises slowly and then more rapidly

as τ increases. An important parameter of this function is the proportionality coefficient in the relationship $G \sim 1/\tau$. Because of this relationship, G can be written in the following form

$$G = D_{RZ} \frac{n_i}{\tau} \tag{69}$$

D_{RZ} = dimensionless factor;
n_i = intrinsic conduction concentration.

For a Shockley-Read recombination center, there is a simple formula for G and also for τ; for τ, however, only when the concentration of recombination centers and the excess concentrations Δn, Δp are very small in comparison with n_D. [72] gives

$$\tau = \frac{1}{C_p} \frac{n_o+n_1}{n_o+p_o} + \frac{1}{C_n} \frac{p_o+p_1}{n_o+p_o} \tag{70}$$

$$G = \frac{C_n C_p n_i^2}{C_n n_1 + C_p p_1} \tag{71}$$

C_n, C_p = recombination rates for electrons and holes, respectively, proportional to the concentration of recombination centers; /100
n_1, p_1 = fictitious carrier concentration, depending on the energy level of the recombination center.

The two preceding equations and Eq. (69) imply

$$D_{RZ} = \frac{n_i}{n_o+p_o} \left(1 + \frac{C_n n_o + C_p p_o}{C_n n_1 + C_p p_1} \right) \tag{72}$$

D_{RZ} obviously depends on the type of recombination center, characterized by C_n/C_p, n_1 and p_1. Therefore, the function $t_d = f(\tau)$ and its existence region, i.e. in particular τ^*, will depend on the type of recombination center. However, detailed computations were carried out only for the case $C_n = C_p$, $n_1 = p_1 = n_i$.

In the studies which were made, it was decided to make the lifetime τ independent of Δn and Δp. This approach represents a great simplification. Strictly speaking, under the conditions of the decay time measurement, neither Δn, Δp, nor the concentration of recombination centers can be neglected. This is one reason that complete quantitative agreement between theory and experiment cannot be anticipated.

6.3. Execution of the Decay Time Measurements

6.3.1. Etching of Edges

In the unetched state, the edge of the p^+np^+ slice -- depending on the history of the slice -- constitutes a short circuit or at least a low-resistance shunt of the p-n junctions. This short circuit or shunt is eliminated by the edge etching. Before the edge etching, the slices are cleaned in hydrofluoric acid and hot aqua regia. For the edge etching, a special teflon apparatus was employed, which exposes only the ring-shaped edge of the slice to the etching attack, while the remainder of the slice is covered and thus protected. CP6[3] was employed as the etchant. The etching lasts for several minutes and depends on the temperature of the etchant and the doping in the vicinity of the border zone. The width of the etched-off edge is about 0.5 mm. The success of the edge etching is assessed by means of the current-voltage characteristic curve. Experience has shown that the assessment is simple, because there are practically no intermediate states between good properties and poor ones. If the edge etching is not successful, it can generally be repeated only once. With further repetitions, the edge of the slice becomes strongly corroded and the blocking properties become poorer instead of better. /101

6.3.2. Electrical Measurement (see Fig. 37)

In the measurement of t_d, the electrodes are a heatable base plate and a top plate pressed onto the p^+np^+ slice by an elastic current lead. (The measuring area contains 20 measuring points.) The voltage u(t) is applied to the capacitor C (variable), which is charged and discharged across the resistance R = 500 Ω by the rectangular pulse generator[4]. The measuring process is periodic with a repetition frequency of 40 Hz, and the duration of the pulse is 8 msec. The time constant RC is small both in comparison with the duration of the pulse and in comparison with the duration of the pauses between the pulses. Therefore, before the end of any given pulse, the voltage u(t) is virtually constant = Û during a period of time on the order of milliseconds, so that the steady state is achieved in the p^+np^+ slice. Moreover, it is ensured that u(t) can decay completely after the end of a pulse before the next pulse starts. The zero-time point (trigger point) for

---

[3] Two parts nitric acid (fuming), one part hydrofluoric acid (40%), one part acetic acid (conc.).

[4] Hewlett Packard, Model 214.

the t_d measurement is at the end of the pulse, i.e. at the beginning of the voltage drop. The current curve i(t) is determined from the voltage curve on the measuring resistance R_M = 100 Ω with an oscillograph.[5]

The series circuit of the p^+np^+ slice and R_M constitutes /102
a highly resistive shunt of the capacitor C to which the voltage u(t) is applied. The current flowing through this shunt is so small that the charging and discharging of C is hardly affected. In particular, this means that u(t) decays very nearly exponentially after the end of the pulse, with the time constant

$$t_a = RC \qquad (73)$$

The voltage on the measuring resistance $R_M i(t)$ is negligibly small in comparison with u(t) during the period of time crucial for the decay time measurement -- from t = -1 msec (about) to $t = t_d$. Therefore, during this period, u(t) is practically equal to the voltage on the p^+np^+ slice. However, the current i(t) depends on both the voltage on the p^+np^+ slice and its derivative with respect to time. This time derivative differs from du/dt by the value $R_M di/dt$. Therefore, there is a falsification of the time derivative of the voltage on the p^+np^+ slice when di/dt assumes large values. This is briefly the case at the beginning of the voltage drop and has the effect that the values of t_d cannot drop below a "blank value." This blank value is a function of R_M and the capacitances of the p-n junctions, and in our measurements it was about 1 μsec.

Transition resistances between the p^+np^+ slice and the electrodes have the same effect as an increased measuring resistance. Isolated occurrences of excessive transition resistances are eliminated by rubbing thin gallium layers onto the p^+np^+ slices with a gallium-wetted aluminum rod.

6.3.3. Polarity and Contacting of $n^+p^+np^+$ Slices

$n^+p^+np^+$ slices with an imbedded n^+ region can be measured in the same fashion as p^+np^+ slices when the polarity and contacting of the slices for the t_d measurement is as depicted in Fig. 40a. /103
The expression "imbedded n^+ region" is to mean that the n^+ region is surrounded by the adjacent region (p^+) on the sides, and is also penetrated by the latter at at least one point (Fig. 40a).

With the given polarity, corresponding to the back direction (negative reverse direction) of thyristors, the following polarities are obtained for the p-n junctions J_1, J_2, and J_o.

[5] Tektronix, Model 543 A with W slide-in unit.

J_1 is polarized in the forward direction, and J_2 in the reverse direction. J_o is short-circuited on the upper side of the slice by the upper electrode, and further inward, where J_o runs parallel to J_1, is weakly polarized in the reverse direction. With this polarity, the imbedded n^+ region acts in essence only as a constriction of the hole-current path from the top of the slice to J_1. Below this constriction, the hole flow lines will spread out. Nevertheless, a certain inhomogeneity of the hole current density will be unavoidable at the level of J_1. However, the conduction mechanism in the t_d measurement is essentially the same as that in a p^+np^+ slice.

If the polarity is different from that given in Fig. 40a, J_2 is polarized in the forward direction, J_1 in the reverse direction, and J_o, in its sections running parallel to J_1, weakly in the forward direction. That is, the n^+ region injects electrons across J_o into the underlying p^+ region from which they can enter the n-type region. The conduction mechanism is then quite different from that in a p^+np^+ slice. We then have a shorted-emitter thyristor polarized in the switch direction, which can under certain circumstances lose its blocking capability (through-connection). Therefore, no measurements were taken out with reversed polarity.

Isolated instances of excessive transition resistances were also found in t_d measurements on $n^+p^+np^+$ slices, and they were eliminated by rubbing on thin gallium layers.

6.4. Comparative Measurements on Diodes

The comparison lifetime measurements were carried out by the retention time method (see Section 2.2.2) on p^+nn^+ diodes prepared from the p^+np^+ or $n^+p^+np^+$ slices following the t_d measurements. /104

6.4.1. Preparation and Measurement

To prepare a p^+nn^+ diode (see Fig. 40b), we removed from the p^+np^+ or $n^+p^+np^+$ slice that p^+ region (by lapping and etching) which had been covered by the positive electrode in the t_d measurement. (That is, with an $n^+p^+np^+$ slice, the imbedded n^+ region was also removed simultaneously.) The p^+n structure obtained in this fashion contains the diffused p-n junction J_2. An antimony-containing gold foil is alloyed into the n-type side of this structure, producing an n-n^+ junction (low-high junction) and the metallic cathode. In the same alloying process, the p^+-type side was connected to the anode (molybdenum) by the Silumin[7] foil, forming an ohmic contact.

[7] Silicon-aluminum eutectic.

For the studies described below, p^+nn^+ diodes were prepared with a silicon-slice diameter (anode diameter) of 7 mm and a cathode diameter of 5 mm. These 7-mm-diameter slices were drilled from larger p^+np^+ or $n^+p^+np^+$ slices with the aid of ultrasonic drills.

Before the lifetime measurement, the diodes were etched for a few seconds in CP6 to obtain a low surface recombination rate. In the measurement of the retention time t_s, the forward current is I_F = 5 mA and the reverse current I_R = 1 mA (cf. Section 2.2.2). From the measured t_s, and the given ratio I_F/I_R, the lifetime τ_K is determined in accordance with Eq. (5). For $I_F/I_R = 5$, (5) implies

$$\tau_K = 1.1 t_s \tag{74}$$

We designate the lifetime here differently than in Sections 1-4, in order to indicate that the objects of measurement differ considerably from each other. The diodes described in Section 2.3 can be termed p^+nR diodes (R = recombination contact), since W_n (width of the n-type region) is large in comparison to L_p (hole diffusion length in the n-type region). Therefore, the slices from which these p^+nR diodes are prepared are sometimes several mm thick. On the other hand, the decay time measurements were conducted on comparatively thin slices (at most 0.4 mm thick), which were largely taken from normal production batches. For the p^+nn^+ diodes prepared from these slices, $W_n \gg L_p$ is in general not satisfied. For short, we call these diodes thin p^+nn^+ diodes. These diodes and also thin p^+nR diodes do not satisfy all the assumptions made for deriving Eq. (5) and thus (74) (see Section 2.1). However, it was found that with a thin n-type region, the lifetime measurement by the retention time method gave better results on p^+nn^+ diodes than on p^+nR diodes. For thin p^+nR diodes, t_s and thus τ_K depend so strongly on W_n, that the influence of lifetime is completely obscured. On the other hand, for thin p^+nn^+ diodes, the influence of the lifetime generally predominates, although here too, W_n depends somewhat on t_s and τ_K. The τ_K value found for a p^+nn^+ diode actually does not agree precisely with the lifetime in the n-type region, but τ_K is a usable measure for the lifetime -- i.e. approximately proportional to lifetime -- whenever the condition $W_n \gtrsim L_p$ is satisfied. Comparing p^+nn^+ diodes with constant W_n, τ_K also increases with the lifetime when $W_n \gtrsim L_p$ is no longer satisfied, but the rise becomes increasingly flatter. /105

Thin p^+nn^+ diodes have the advantage over thin p^+nR diodes in lifetime measurement, but they have the drawback that a heat treatment (alloying process) must be carried out to produce the n-n^+ junction. In this process, small quantities of additional

recombination centers enter the n-type region. The lifetime in the n-type region of the p^+nn^+ diode is therefore somewhat lower than in the n-type region of the original p^+np^+ or $n^+p^+np^+$ slice. At the relatively low alloying temperature, this side effect is usually of relatively minor importance and will henceforth be ignored. /106

The n-type regions of the p^+nn^+ diodes prepared for the studies now to be described were always about 40 μm thinner than the original p^+np^+ or $n^+p^+np^+$ slices.

6.4.2. Relationship Between the Measured Lifetime τ_K and the Recovery Time

Since on one hand, the value of τ_K measured for p^+nn^+ diodes is a measure of the lifetime, and since on the other hand, the recovery time[8] t_q of a thyristor is a linear or at least monotonically increasing function of lifetime, t_q should be a linear or at least monotonically increasing function of τ_K. In practice, the relationship between t_q and τ_K can be checked as follows. From batches of p^+np^+ or $n^+p^+np^+$ slices of a specific type, several specimens are taken per batch to prepare p^+nn^+ diodes and τ_K is measured for them; t_q is measured on the thyristors produced from the main portion of each batch. The mean values $\overline{t_q}$ and $\overline{\tau_K}$ are plotted against each other for each batch. Taking mean values is necessary because t_q and τ_K sometimes vary widely within a batch. The specimens for preparing p^+nn^+ diodes are taken after the last heat treatment having a substantial effect on lifetime (gettering process, gold diffusion). With some types of thyristors, this heat treatment is followed by an alloying process. This is particularly true for diffused-alloyed thyristors, where the diffused p^+np^+ structure is converted to an $n^+p^+np^+$ structure by alloying on an antimony-containing gold foil. The alloying process for thyristors resembles the alloying process for p^+nn^+ diodes (Section 6.4.1) and therefore has the same, usually negligible influence on lifetime.

Figure 41 depicts the relationship between t_q and τ_K for a diffused-alloyed thyristor type. Each batch supplies one /107

[8] The recovery time t_q is the minimum waiting time between the zero-current point in passing from the switch direction to the back direction and the return of a voltage in the forward direction which must pass until the thyristor regains its blocking capability in the forward direction. t_q is a function of, among other things, the temperature of the thyristor slice and the preceding forward current [73, 74].

measurement point. The p^+np^+ slices of a single batch all come from the same crystal, were gallium-diffused together in an ampule, and were gold-diffused at the same temperature -- whenever a gold diffusion was carried out. The slices of different batches came from different crystals, but are all of the same type. Six slices were taken for the τ_K measurement from each batch, and t_q was measured for about 50 thyristors. There is a linear relationship between $\overline{t_q}$ and $\overline{\tau_K}$, but the measured points fluctuate in spite of the averaging. As the result shows, it is still possible to predict with some certainty the mean recovery time $\overline{t_q}$ by measuring τ_K on a relatively small number of specimens per batch. If $\overline{t_q}$ lies outside the desired range, the lifetime can be raised or lowered by gettering or gold diffusion.

A direct check on the relationship between t_q and τ_K can be made in the following manner. After the t_q measurement, the $n^+p^+np^+$ slices were prepared from -- e.g. fully diffused -- thyristors, and p^+nn^+ diodes were produced from these slices and τ_K measured for these diodes. Figure 42 shows the result of such a series of measurements on fully diffused thyristors. Three diodes (7 mm diameter) were produced from each prepared $n^+p^+np^+$ slice. τ_K is averaged over the three diodes per slice. In spite of occasional large fluctuations, it can be seen from Fig. 42 that t_q increases with increasing τ_K. The fluctuations are due at least in part to the inhomogeneity of the $n^+p^+np^+$ slices relative to lifetime. A clear sign of such inhomogeneities is provided by the fluctuations in τ_K values for the three diodes from a single slice: for some of the slices, the span[9] of the τ_K values was at least 30% of the mean τ_K.

6.5. Results and Expansion of the t_d Measurement

Slices (p^+np^+, $n^+p^+np^+$) and fully diffused thyristors (functional components, but without case) were studied. The t_d values of the slices were correlated with the τ_K values of the p^+nn^+ diodes prepared from them. The t_d values of the thyristors were compared directly with their recovery times. In the decay time measurements, the "plateau voltage" $\hat{u}$ (see Fig. 37) was always 50 V. The time constant t_a of the voltage drop was sometimes constant = 100 µsec, and sometimes, in elaboration of the method, t_a was adjusted "individually" in accordance with specific criteria for each test object (Section 6.5.2). /108

6.5.1. Gallium-Diffused p^+np^+ Slices

The slices studied, all of the same type, came from seven diffusion batches and 14 crystals. After the gallium diffusion,

[9] Difference between the largest and smallest of the three τ_K values for a single slice.

the p^+np^+ slices underwent no further heat treatment (gettering, gold diffusion). The type of slice is characterized by the following figures:

n-type region:	ρ_n = 65 Ωcm, W_n = about 210 μm
p^+-type regions:	N_S(Ga) = about $5 \cdot 10^{18}$ cm^{-3}, thickness = = about 95 μm
diameter	= 14 mm (before edge etching, afterward about 13 mm)

Here:

ρ_N = resistivity of n-type region,
W_n = width (thickness) of n-type region,
N_S(Ga) = surface concentration of gallium.

The decay time measurements were carried out with t_a = 100 μsec at ϑ = 130°C, where the critical temperature $\vartheta(t_a = 100\ \mu sec)$ was between 136 and 148°C. A p^+nn^+ diode (7 mm diameter) was prepared from each p^+np^+ slice. τ_K was measured at 130 and 25°C.

In diagrams A and B of Fig. 43, the measured lifetimes τ_K(130°C) and τ_K(25°C) are plotted against t_d(130°C). In spite of considerable fluctuations, it can be seen in both diagrams that τ_K is approximately a linear function of t_d. Diagram A is the more "relevant" of the two diagrams, because there the measurement points were obtained by measuring t_d and τ_K at the same temperature. Since lifetime is a function of temperature, it initially appears doubtful that there is a relationship between τ_K(25°C) and t_d(130°C). (Measuring t_d at 25°C would be pointless because t_d depends on lifetime only in a relatively narrow range of elevated temperatures, cf. Section 6.2.3 and [71]. However, it was found that in going from 130°C to 25°C, τ_K changes by a factor which is roughly the same for all diodes; roughly $\tau_K(25°C)/\tau_K(130°C) = 0.3$. Hence, diagram B yields roughly the same picture as diagram A. /109

The scattering of the measurement points in diagram A cannot be explained by the measurement errors alone, which can be about ±2% in the τ_K measurement and about ±5% in the t_d measurement. Other causes for the scattering could be:

a) In measuring t_d, the surface current flowing across the edge of the slice might not be negligible in comparison with the volume current, even when the edge etching appears successful. The surface current will not have the same time behavior as the volume current, and therefore falsify the decay time, which is a volume effect. The falsification of the value of t_d can vary from one slice to the next depending on the ratio of surface and volume currents and on the time behaviors of the two current components.

b) Inhomogeneities of lifetime in the lateral direction (parallel to the p-n junctions) will not have the same effect on t_d as on τ_K. This is particularly true when, as in the present case, the p^+nn^+ diode has a smaller diameter than the original slice on which the t_d measurement was carried out.

c) The relationship between decay time and lifetime depends on the type of recombination center (Section 6.2.3.2). A similar effect can be anticipated when the slices contain several types of recombination centers, in proportions varying from one slice to the next. This will produce scattering in the τ_K-t_d diagram, since the function $\tau_K = F(t_d)$ will depend on these proportions.

In passing from diagram A to diagram B, there is an additional cause for the scattering: the ratio $\tau_K(25°C)/\tau_K(130°C)$ is not precisely constant, but generally differs somewhat for different diodes. /110

The τ_K-t_d relationships considered in the following sections are depicted as in diagram B of Fig. 43: $\tau_K(25°C)$ is plotted against the decay time, which was measured at a suitably high temperature (130 or 140°C). In this representation as well, the τ_K-t_d relationship still depends on the type of slice.

6.5.2. Expansion of the Decay Time Measurement

In earlier measurements of t_d [71] and in those described in the preceding Section 6.5.1, the time constant of the voltage drop had a fixed value: t_a = 100 µsec. The procedure for measuring t_d with t_a = const. has the drawback that the condition $\theta < \Theta$ (t_d-measuring temperature smaller than the critical temperature) must be satisfied so that a t_d value will exist at all. With fixed t_a and Q, Θ depends on, among other things, ρ_n (resistivity of the n-type region), and in fact Θ increases as ρ_n decreases. The closer θ is to Θ, the clearer the relationship between τ_K and t_d, so that matching θ to the ρ_n of the slice type cannot be completely avoided. The scattering of Θ within a single slice type causes even more problems than the ρ_n-dependence of Θ. Namely, the range of variation of Θ will be known to a certain extent only when Θ has been determined for a large number of slices.

In order to avoid the drawbacks associated with measuring t_d (t_a = const), the t_d measurement was expanded by two modifications in each of which the condition t_a = const. was replaced by a different one. The two modifications have in common that the current-time curve is in effect forced to have a minimum -- i.e. t_d will exist -- where t_a is set "individually" for each test object. This simplifies the choice of the measuring temperature θ.

If for a given slice the time constant t_a of the voltage drop is varied, while all other parameters are held constant, the form of the current-time curve during the voltage drop changes. If t_a is sufficiently small, the current-time curve has a minimum, which -- as opposed to that depicted in Fig. 37 -- lies under the zero line. I.e. t_d exists and $i_d < 0$, where i_d designates the current at the time t_d, i.e. at the minimum. As t_a increases, so do both t_d and i_d. At a certain $t_a = t_{ao}$, $i_d = 0$, i.e. the current minimum lies exactly on the zero line (see Fig. 44a). As t_a continues to increase, so do t_d and i_d; i_d is now positive, and the case depicted in Fig. 37 is obtained. When t_a reaches a certain value t_{ap}, the current-time curve no longer has a minimum, but instead a point of inflection with a horizontal tangent (see Fig. 44b). For $t_a > t_{ap}$, i(t) is a monotonically decreasing function which finally intersects the zero line. /111

This t_a-dependence of the i(t) curve, which can be explained theoretically, contains two characteristic cases, $i_d = 0$ and $t_a = t_{ap}$, from which two new conditions for measuring t_d can be established.

The first condition reads: t_a is to be adjusted so that $t_a = t_{ao}$, i.e. $i_d = 0$ is satisfied. The resulting decay time -- which is sure to exist because of $t_{ao} < t_{ap}$ -- will henceforth be designated $t_d(i_d = 0)$ (see Fig. 44a).

The second condition is as follows: t_a is first adjusted so that $t_a = t_{ap}$ is satisfied. Then t_a is halved. The decay time measured at that point -- which is sure to exist because of $t_a = t_{ap}/2$ -- is designated $t_d(t_{ap}/2)$ (see Fig. 44b).

The required variation of t_a is accomplished by adjusting the variable[10] capacitance C with constant observation of the i(t) curve. The values of t_{ao} and t_{ap} are generally different for different slices. /112

The relationship between the lifetime and the two variants of the decay time has not been subjected to a more precise theoretical investigation. From more qualitative considerations, it can be inferred that both $t_d(i_d = 0)$ and $t_d(t_{ap}/2)$ increase as the lifetime increases, if, with constant Q and θ, slices of a specific type, containing the same type of recombination centers, are compared with each other.

---

[10] C = 0.001-10 μF (4 decades). In the measurements, t_a was never larger than 1 msec and was therefore small in comparison with the pulse duration (8 msec) and the duration of the pauses between the pulses (17 msec), cf. Section 6.3.2.

While $t_d(t_a = \text{const})$ rises rapidly with increasing temperature θ up to the critical temperature Θ, $t_d(i_d = 0)$ decreases with increasing θ. The temperature variation of $t_d(t_{ap}/2)$ is comparatively weak, but at high θ $t_d(t_{ap}/2)$ also falls with increasing θ. Hence, θ cannot be chosen arbitrarily large for either $t_d(i_d = 0)$ or for $t_d(t_{ap}/2)$; it is true that there is no critical temperature for the two kinds of t_d, but if θ gets too large, the measured values become very small and no longer provide any information on the lifetime. There is also a lower limit for temperatures suitable for measuring the two types of t_d. Namely, if θ is too small, the precision of measurement becomes very poor because of the smallness of i_o, and t_a must be set to high values which are no longer small relative to the duration of the generator pulse (see Section 6.3.2). Hence, for both $t_d(i_d = 0)$ and $t_d(t_{ap}/2)$, there are only bounded suitable temperature ranges in which t_d can be measured with sufficient precision and in which a relationship with lifetime can be anticipated. However, these temperature ranges are wider than the temperature range suitable for measuring $t_d(t_a = 100\ \mu\text{sec})$. Frequently, the three temperature ranges overlap, so that all three decay times can be measured at the same temperature.

6.5.3. Gallium-Diffused, Gettered, and Gold-Diffused p^+np^+ Slices

The slices came from one crystal and were all of the same type:

n-type region:	ρ_n = about 70 Ωcm, W_n = about 210 μm
p^+-type regions:	N_S(Ga) = about $5 \cdot 10^{18}$ cm^{-3}, thickness = = about 95 μm
diameter	= 19 mm (before edge etching, afterward about 18 mm).

/113

Six groups of such slices were studied:

A = only gallium-diffused;
B = gallium-diffused and gettered;
C = gallium-diffused and strongly gettered;
D_1 = gallium diffused, strongly gettered, and gold-diffused for 1 hour at 780°C
D_2 = gallium-diffused, strongly gettered, and gold-diffused for 1 hour at 800°C
D_3 = gallium-diffused, strongly gettered, and gold-diffused for 1 hour at 830°C.

For the p^+np^+ slices, $t_d(t_a = 100\ \mu\text{sec})$, $t_d(i_d = 0)$, the critical temperature $\Theta(t_a = 100\ \mu\text{sec})$, and the steady-state reverse current i_o were measured. Two or three p^+nn^+ diodes per slice were prepared for the τ_K measurement.

The most important results are collected in Table 10. For each quantity measured, the table gives the smallest and largest values as well as the mean (underlined) within a group; τ_K is already averaged over the two or three diodes from one slice.

First considering τ_K, the expected trend is found from Group A to D_3: gettering increases lifetime, gold diffusion reduces it -- the higher the gold diffusion temperature, the greater the reduction period. The lifetimes in Group A are strikingly short, in general the lifetimes for gallium-diffused p^+np^+ slices (only gallium diffusion) are longer.

$t_d(t_a = 100\ \mu sec)$, measured at 130°C, shows a variation from Group A to D_1 similar to that of τ_K. Measurements cannot be taken for Groups D_2 and D_3, because the critical temperature Θ is below 130°C. The drop in Θ with increasing gold diffusion temperature cannot be explained by the ρ_n-dependence of Θ mentioned in Section 6.5.2, because gold diffusions resulted in only small changes of ρ_n. The drop in Θ is probably based on the properties of gold as a recombination center, which we will discuss further below. It is practically impossible to measure $t_d(t_a = \text{const.})$ for gold-diffused and non-gold-diffused slices at the same temperature. At a lower temperature, e.g. 100°C, $t_d(t_a = 100\ \mu sec)$ would be measurable for the gold-diffused slices as well. However, then the t_d values for non-gold-diffused slices would be very low and would not exhibit any relationship to lifetime. /114

$t_d(i_d = 0)$, likewise measured at 130°C, can also be measured for the strongly gold-diffused Groups D_2 and D_3, since after all there is no critical temperature for the t_d variants. Like τ_K, $t_d(i_d = 0)$ is increased by gettering (from Group A to C) and reduced by gold diffusion (from Group C to D_3). If τ_K is plotted against $t_d(i_d = 0)$ for slices of short lifetime (Group A, D_1, to D_3), two features are worth noticing (Fig. 45a). First, the measurement points of Group A with the low "natural lifetime" (after gallium diffusion) are far below the curve which can be drawn through the measurement points of the gold-diffused groups. Second, the values of $t_d(i_d = 0)$ for the gold-diffused slices are very small (1 µsec) even at relatively large values of τ_K, i.e. the blank value of the measuring circuit (about 1 µsec) is practically reached. This implies that for gold-diffused slices with lifetimes below a certain limit, even $t_d(i_d = 0)$ fails as a measurement of lifetime.

In considering the steady-state reverse current i_o at 130°C (Table 10), a conspicuous feature is the large difference between the two Groups A and D_3 with short lifetimes. Moreover, i_o falls with increasing τ_K for the gold-diffused slices, while i_o rises with increasing τ_K for the non-gold-diffused slices

(nevertheless, i_o rises only by a factor of 2, while τ_K is growing by a factor of 30).

Gold-diffused slices with short lifetimes (Group D_3) thus /115
differ greatly from slices with low "natural lifetimes" (Group A) in both critical temperature $\Theta(t_a = \text{const.})$ and steady-state reverse current i_o. These differences are comprehensible qualitatively, given the two following assumptions. First: the D_3 slices essentially contain only gold atoms as recombination centers, surely a plausible assumption in view of the long lifetime before gold diffusion. Second: the A slices contain no gold, but another, unknown recombination center, whose energy level is at least 0.18 eV away from the center of the forbidden band. It can then be shown that the coefficient D_{RZ} introduced in Section 6.2.3.2 (Eqs. (69) and (72)) is larger for the D_3 slices than for the A slices when the values from [24] and [75] are employed for the energy levels and recombination rates of gold. For a given type of slice and for given u, t_a, and θ, the behavior and the existence region of the function $t_d(t_a = \text{const.}) = f(\tau)$ depend on D_{RZ} (i.e. on the type of recombination center). In the range $0 < \tau < \tau^*$, $t_d(t_a = \text{const.})$ does not exist (cf. Section 6.2.3.2). This means: for $\tau < \tau^*$, the measuring temperature $\theta >$ the critical temperature Θ; for $\tau > \tau^*$, however, $\theta < \Theta$, i.e. $t_d(t_a = \text{const.})$ exists. With the aid of the formulas collected in [71], it can be shown that τ^* increases with increasing D_{RZ}. This is plausible because a larger D_{RZ} means a larger generation rate G; hence, not until large τ will the generation of charge carriers in the space charge region be small enough so that i(t) will pass through a minimum. Since τ^* grows with increasing D_{RZ}, and since the D_{RZ} of the A slices is smaller than that of the D_3 slices, it follows that in a certain lifetime region the decay time $t_d(t_a = \text{const.})$ can be measured for the A slices or slices "lifetime-doped" with the same recombination center, but not for slices "lifetime-doped" with gold. No quantitative calculation of this lifetime range was attempted. The experimental results indicate that the upper /116
limit of this range is $\tau_K = 16$ µsec, where the $\Theta(t_a = 100$ µsec)-τ_K curve of the gold-diffused slices passes through 130°C. Nevertheless, it must also be assumed that gold is the essential lifetime doping in all three groups D_1, D_2, and D_3, but this assumption appears risky for Group D_1. The lower limit of the lifetime region defined above is below $\tau_K = 1.7$ µsec, since $t_d(t_a = 100$ µsec) is still measurable even for the A slices.

From Fig. 45b, it can be deduced that the steady-state reverse current i_o of the A slices is smaller than that of gold-diffused slices with the same lifetime. This can be explained with the same assumptions made above to interpret the different behaviors of Θ as τ_K varied. Namely, i_o is a

function of T_{RZ}[11], and in fact i_o rises as D_{RZ} increases. Furthermore, theory[11] indicates that the reverse current i_o increases as lifetime decreases if D_{RZ} is constant. The variation of τ_K with $1/i_o$ (Fig. 45b) found for the gold-diffused slices is therefore in essence comprehensible theoretically. No explanation has yet been found for the rise in i_o with increasing τ_K in the non-gold-diffused slices (see Table 10, Groups A-C).

Although the relationship between $t_d(i_d = 0)$ and the lifetime has not yet been subjected to a more thorough theoretical study, it can be assumed that this relationship as well will depend on D_{RZ}, i.e. on the type of recombination center. With this assumption, it is understandable that the measured points of Group A in Fig. 45a lie off the curve through the measurement points of the gold-diffused slices.

6.5.4. $n^+p^+np^+$ Slices

In one series of measurements, $t_d(t_a = 100\ \mu sec)$ and $t_d(i_d = 0)$ were compared as to their suitability for lifetime measurements on $n^+p^+np^+$ slices. The $n^+p^+np^+$ slices not lifetime-doped with gold came from a large number of crystals and diffusion batches and exhibited the following parameters characterizing the type: /117

n-type region:	ρ_n = about 30 Ωcm, W_n = about 150 μm
p^+-type regions:	N_S = about $5 \cdot 10^{18}$ cm^{-3}, thickness = about 50 μm
n^+-type region:	N_S = about 10^{21} cm^{-3}, thickness = about 25 μm (imbedded in a p^+-type region)
diameter	= 19 mm (before edge etching)

Two or three p^+nn^+ diodes per slice were prepared for measuring τ_K.

In Fig. 46a, τ_K, measured at 25°C and averaged over the diodes of one slice, is plotted against $t_d(t_a = 100\ \mu sec)$, measured at 140°C. In Fig. 46b, $\tau_K(25°C)$ is plotted against $t_d(i_d = 0, 140°C)$ for the same slices. Comparing the two diagrams shows that $t_d(i_d = 0)$ is far more suitable than $t_d(t_a = 100\ \mu sec)$ as a measure for the lifetime. In Fig. 46a, the scattering of the measurements above $t_d(t_a = 100\ \mu sec) = 30\ \mu sec$ is so great that drawing a curve through the measurements would be pointless. I.e. $t_d(t_a = 100\ \mu sec)$ is practically unusable as a measure for lifetime. On the other hand, Fig. 46b shows a linear relationship between τ_K and $t_d(i_d = 0)$, in spite of considerable scattering in this case as well.

[11] Cf. [71], Eqs. (20) and (30).

The drawback of the "primary" decay time $t_d(t_a = 100\ \mu sec)$ found in this series of measurements is not caused by too low a measuring temperature ($\theta = 140°C$). Namely, the lower limit of the range of the critical temperature Θ was 146°C. Increasing θ, which might have resulted in a better relationship between $t_d(t_a = 100\ \mu sec)$ and τ_K, was therefore not possible in practice.

No satisfactory explanation for the failure of $t_d(t_a = 100\ \mu sec)$ as a measure for the lifetime of the $n^+p^+np^+$ slices has yet been found. In particular, the influence of the imbedded n^+-type region is unclear.

6.5.5. Relationship Between Decay Time and Recovery Time for Fully Diffused Thyristors

/118

Up til now, the utility of the decay time method has been studied by investigating the relationship between t_d and the lifetime τ_K measured by another technique. In the practice of component production, one is less interested in the relationship between the lifetimes measured by different techniques than in the relationship between lifetime measurements and the electrical properties of the component, e.g. the recovery time t_q [73, 74]. That there is a relationship between τ_K and t_q was shown in Section 6.4.2. In the direct examination of the τ_K-t_q relationship -- measurement of both quantities on the same object -- the thyristor had to be destroyed to prepare p^+nn^+ diodes for measuring τ_K. In contrast, a direct examination of the t_d-t_q relationship can be carried out nondestructively.

Functional, fully diffused thyristor elements (with massive anode and cathode, but without case) were investigated with slices of the following type:

n-type region:	ρ_n = about 35 Ωcm,	W_n = about 150 µm
p^+-type regions:	$N_S = 5 \cdot 10^{18}\ cm^{-3}$,	thickness = about 65 µm
n^+-type region:	N_S = about $10^{21}\ cm^{-3}$, (imbedded in a p^+-type region)	thickness = about 30 µm

slice diameter = 23 mm

The thyristor slices came from several crystals and diffusion batches (gallium, phosphorus). In order to lower the lifetime, some of the slices were also gold-diffused at temperatures between 835 and 845°C.

In the t_d measurements, the cathode was positive and the anode was negative. The p-n junction between the n^+-type region and the adjacent p^+-type region is largely short-circuited by the cathode. Therefore, the practical situation is that depicted in Fig. 40a.

The recovery time t_q is plotted against $t_d(i_d = 0)$ and $t_d(t_{ap}/2)$ in Fig. 47. t_q was measured at 125°C and a forward current of 100 A, the two kinds of t_d at 130°C. /119

For the non-gold-diffused thyristors, a clear relationship is found between t_q and $t_d(i_d = 0)$ as well as $t_d(t_{ap}/2)$, even though accompanied by the usual scattering. On the other hand, both t_d variants fail for the gold-diffused thyristors, whose recovery times are mostly less than 30 µsec, thus meeting the requirements for fast thyristors. The measurements for the gold-diffused thyristors (19 items) lie in the shaded regions, i.e. far from the other measurements. For the gold-diffused thyristors, no variation of t_q with t_d is detectable. Moreover, most of the values of t_d -- both at $i_d = 0$ and at $t_a = t_{ap}/2$ -- are less than 3 µsec, and thus close to the "blank value" of the circuit, which is about 1 µsec (cf. Sections 6.3.2 and 6.5.3).

As mentioned in Section 6.5.2, there are for $t_d(i_d = 0)$ and $t_d(t_{ap}/2)$ only limited temperature ranges suitable for lifetime determination. For the non-gold-diffused thyristors, even those with low recovery times, i.e. with low lifetimes, $\theta = 130°C$ is in the suitable temperature ranges. On the other hand, for gold-diffused thyristors, $\theta = 130°C$ is too high a measuring temperature, which is indicated by the very low t_d values. If there are suitable temperature ranges at all for the gold-diffused thyristors, in which a relationship can be found between lifetime or recovery time on the one hand and $t_d(i_d = 0)$ or $t_d(t_{ap}/2)$ on the other, then these ranges are below $\theta = 130°C$. However, measurements at lower temperatures -- down to $\theta = 80°C$, at which t_d measurements were still just possible -- led to disappointing results. Indeed, the t_d values increased with decreasing θ, but no t_q-t_d relationship could be distinguished because of the large fluctuations in the measurements.

The difference in behavior in decay time measurements between gold-diffused thyristor slices with short lifetimes on the one hand and non-gold-diffused thyristor slices with short lifetimes on the other corresponds completely to the difference in behavior described in Section 6.5.3 between gold-diffused and non-gold-diffused p^+np^+ slices. Accordingly, it can be assumed that a recombination center other than gold predominates in non-gold-diffused thyristor slices. /120

Measuring t_d with constant $t_a = 100$ µsec at $\theta = 130$ and 140°C in this sequence of measurements did not produce any useful result either. Measurements could not be taken for the gold-diffused thyristors -- the critical temperatures were far under 130°C -- which is not surprising in view of the results of Section 6.5.3. If t_q is plotted against $t_d(t_a = 100$ µsec) for the non-gold-diffused thyristors, a picture is obtained

which is very similar to that of Fig. 46a. Therefore, we must discard $t_d(t_a = 100\ \mu sec)$ as a measure for the recovery time.

According to the results of this and the preceding section 6.5.4, the t_d values determined under the conditions $i_d = 0$ and $t_a = t_{ap}/2$ are a useful measure for the lifetime or recovery time of non-gold-diffused $n^+p^+np^+$ slices or thyristors. On the other hand, both types of t_d fail when gold diffusion has depressed the lifetime to such low values that the low recovery times required for fast thyristors are obtained. The "primary" decay time $t_d(t_a = 100\ \mu sec)$ did not work for either gold-diffused or -- for unexplained reasons -- non-gold-diffused $n^+p^+np^+$ slices and thyristors.

6.5.6. t_d Measurements on p^+np^+ Slices for Hall Effect Studies

In Section 5.3.5, it was shown that there is a relationship between t_d or $1/t_d$ and the concentration of deep impurities in the n-type region determined by Hall measurements (see Figs. 30 and 32). The t_d values were measured under the condition $t_a = t_{ap}/2$ at $\theta = 130°C$: $t_d(t_{ap}/2, 130°C)$.

The gallium-diffused p^+np^+ slices came from four crystals, which differed considerably from one another in resistivity (from about 70 to about 150 Ωcm). The ρ_n of the diffused slices /121 differed to a greater or lesser degree from the initial values of the crystals. For decay time measurements on slices with different ρ_n (resistivity of n-type region), the most suitable value is $t_d(t_{ap}/2)$. This is because this measurement can be used as a measure for lifetime in a relatively broad temperature range and depends comparatively weakly on temperature, and the position of this temperature range depends only weakly on ρ_n.

6.6. Summary

The decay time method is based on the following experimental finding, which can be interpreted theoretically to a great extent: If an initially constant, and then (from $t = 0$) exponentially decaying (falling) voltage is applied to a p^+np^+ slice held at elevated temperature, under certain conditions the current will pass through a minimum; the time at this minimum, called the decay time t_d, proves to be a measure of the lifetime (of the n-type region) of the p^+np^+ slice in many cases.

To test the method, a large number of diffused p^+np^+ and $n^+p^+np^+$ slices (with imbedded n-type region) were studied. In this case, t_d was compared with the lifetime τ_K measured by the retention time method. τ_K was measured on p^+nn^+ diodes prepared from the p^+np^+ and $n^+p^+np^+$ slices after the measurement of t_d.

The decay time measurements were first taken with a fixed time constant t_a for the voltage drop: t_a = 100 µsec. The decay time $t_d(t_a$ = 100 µsec) increases with increasing temperature, but only exists up to a "critical temperature" Θ, which depends on, among other things, the resistivity of the n-type region. There can be a relationship with lifetime only in a relatively narrow temperature range below Θ.

For gallium-diffused p^+np^+ slices, an approximately linear relationship was found between τ_K and $t_d(t_a$ = 100 µsec). In this case, the measuring temperature was close (about 10°) to the lower limit of the range of variation of Θ. This Θ range was about 10° for gallium-diffused slices of a specific type. Slices with short lifetimes after gallium diffusion and slices with longer lifetimes due to gettering also lie in this range. However, gallium-diffused slices whose lifetimes had been reduced from high values to low ones by gold diffusion exhibited a strong drop in Θ with decreasing τ_K, which made the decay time measurement with t_a = constant practically impossible. /122

For $n^+p^+np^+$ structures, $t_d(t_a$ = 100 µsec) proved to be unsuitable for lifetime measurements. This is because there was no sufficiently well-defined relationship between decay time and lifetime for non-gold-diffused structures. For gold-diffused structures, the same difficulties were encountered as with gold-diffused p^+np^+ structures.

The decay time method was expanded to two modifications in which the time constant t_a is set individually for each slice according to specific criteria. There is no critical temperature for the two t_d variants. Here too, however, a relationship with lifetime will exist only in limited, although relatively broad temperature ranges.

For non-gold-diffused $n^+p^+np^+$ slices and thyristors, approximately linear relationships were found with τ_K and with the recovery time t_q for both t_d variants down to low values of τ_K and t_q. However, if short recovery times were produced by gold diffusion, then both t_d variants failed as a measure of recovery time and lifetime.

Since slices with low lifetimes without gold diffusion behave completely differently in the decay time measurement from gold-diffused slices with low lifetimes, it is assumed that a recombination center other than gold dominates in non-gold-diffused slices. According to a rough estimate, the energy level of this recombination center must be at least 0.18 eV away from the center of the forbidden band. /123

7. Application of Gold Diffusion to the Production of Fast Diodes /124

7.1. Introduction

For many applications in power electronics, components such as diodes and thyristors with particularly short turn-off times are desirable. By systematically controlling the carrier lifetime, these components can be adapted to the intended utilization.

A method must be found to reproducibly diminish the lifetime by systematic injection of recombination centers.

In the following, we will describe such experiments in the development of a fast power diode. First it was necessary to study the relationships between the lifetime and the various parameters of the treatment.

7.2. Program

The customary method for reducing lifetime is the controlled injection of gold into the silicon lattice. This process depends on the amount of gold applied to the slice, the temperature at which the gold is diffused into the crystal, and the diffusion time. We will investigate the influence of these three parameters on the carrier lifetime.

Moreover, we also wish to establish the extent to which the type of gold application influences the result. The work will be systematically directed toward the practical application, in order to develop a fast power diode. For thyristors, the control of specific parameters such as recovery time by gold diffusion is associated with greater difficulties than with diodes. The differences in technological procedure between diodes on the one hand and thyristors on the other will be discussed.

7.3. Preparation and Measuring Instruments /125

Gallium-diffused p-n-p slices with the longest possible lifetimes were employed for the investigations. The slices were taken from batches routinely diffused by uniform procedures. As far as gallium penetration depth and gallium surface concentration, the differences between the diffusion batches were small.

The gallium-diffused slices were coated with gold. In the majority of cases, this took place in an aqueous solution of chloroauric acid under the action of ultrasonics. After drying, the slices were placed in a three-zone tube furnace, where they

were diffused under flowing hydrogen (about 150 ℓ/hour). The temperature of the furnace was held constant within ±1°C. After diffusion, the slices were lapped on one side (a p-n structure was produced from the p-n-p structure), etched, and contacts alloyed on (cathode side Au(Sb), anode side aluminum-silicon eutectic on molybdenum carrier plate). After a further etching, the carrier lifetime of the finished diode was measured. A Tektronix oscillograph with a slide-in S-unit was employed for the lifetime measurement. This unit switches the current through the diode from a specific forward current (e.g. 5 mA) to the value zero or to a specific reverse current (e.g. 1 mA), so that lifetime can be measured in accordance with Lederhandler and Giacoletto [7] and Kingston [8].

7.4. Gold Coating

We attempted to discover whether a different coating of the silicon slices with gold would have an influence on the lifetime obtained. For this purpose, three identical groups of Si slices (30 mm diameter) were coated with gold in different ways, and then identically treated and measured together.

The first group was coated in aqueous gold solution (0.001% solution of chloroauric acid) under the action of ultrasonics and then dried.

The second group received a thin gold coating (from the vapor; about 0.1 μm thick) on one side and the third group only on one half of one side of the slice. /126

After the joint diffusion of the three groups, a region of the slices was lapped off and test wafers 7 mm in diameter were drilled out. In the third group, we made sure that wafers were also taken from the half of the slice not coated with gold.

Analysis yielded no difference in lifetimes between the three groups. All values of τ_K (lifetime measured by the Kingston [8] method) were between 1.9 and 1.9 μsec after diffusion for 1 hour at 845°C. Neither could any difference be established between the wafers obtained from the coated and uncoated halves of the third group.

This result can be explained by the known high volume diffusion rate [24] of gold in silicon. Even with nonuniform application, the gold spreads throughout the silicon slice in a relatively short time.

Based on these results, the simplest procedure of gold coating, namely application in aqueous solution, was always used in further experiments. The slices were consistently treated for 30 minutes in the gold solution with ultrasonics.

In order to hold the amount of gold precipitated on a unit of area as constant as possible, a fixed volume of solution per unit of slice area was always employed, and the concentration of the solution always held constant.

7.5. Dependence of Lifetime on the Diffusion Parameters Time and Temperature

Once the gold coating had been applied, the slices were diffused for different periods of time at specific temperatures, and lifetimes measured on the test diodes. Figure 48 shows lifetime as a function of temperature and time. Both the Kingston value τ_K and the Giacoletto value τ_e are plotted. Above 770°C, an almost exponential drop in lifetime with increasing gold diffusion temperature can be observed. Nevertheless, with increasing temperature, there is a considerable rise in the resistivity ρ, which begins as early as 800°C, as depicted in Fig. 49. The dependence of lifetime on the diffusion time is very small. Hence, the logical way to attain a specific τ_K is by regulating the temperature. /127

In order to discover whether limiting the supply of gold to the surface of the slice, i.e. utilizing a gold source with a finite capacity, has any influence on the lifetime obtained, further experiments were conducted.

The p-n-p slices were first coated with gold as usual in the aqueous solution. Some of the slices were gold-diffused at different temperatures (790-840°C) and times (0.5, 1, 2, and 3 hours) by the previously described procedure. After the application of the gold coating, the remainder of the slices were brought to the diffusion temperature for 2 minutes and then cooled. Next the gold was chemically removed and the diffusion continued for the full time (30 or 60 min) at the previously chosen temperature (interrupted diffusion). The result is depicted in Fig. 50. It can be seen that the lifetime values from the interrupted diffusion are markedly higher than those from normal diffusion. Here the limited capacity of the gold source obviously plays a role. Only the gold atoms injected into the silicon lattice during the short preliminary diffusion of 2 minutes are available for the longer principal diffusion.

In tests on p-n-p-n^+ slices from thyristor production, curves were obtained corresponding precisely to those in Figs. 48 and 50.

The diffusion time was prolonged to as much as 6 hours for individual slices. The resulting lifetimes did not differ much from those obtained with a 3-hour diffusion time. In the practical application, the desired shortening of the lifetime is concluded within a relatively short time (1 hour).

7.6. Development of a Fast Power Diode /128

For many applications, diodes are required which transpose rapidly from the conducting to the blocking state (fast recovery diode). For this a short carrier lifetime is necessary and is achieved with gold diffusion. In 83 diffusions, the influence of diffusion temperature on lifetime and thus on the quantities crucial for the operation of the component such as the forward voltage drop U_F and the storage charge Q_{STG} (Fig. 51) was studied. Temperatures in the range from 840°C to 924°C were employed with diffusion times of 1 hour in each case. The starting materials were normal gallium-diffused p-n-p slices for the D 300 diode (slice diameter 19 mm, thickness of n-type base d_n = 150 µm), since the diode to be developed was intended to have a continuous limiting current of about 250 A (D 300 = = diode for 300 A continuous limiting current). In Fig. 52, the values of U_F and Q_{STG} obtained from various gold diffusions are plotted, giving the scattering at 880°C. The two curves run in opposite directions. The electrical-engineering requirements on the diodes place limits on U_F and Q_{STG} (Q_{STG} and U_F should not exceed a maximum value). There is only a narrow range in which both U_F and Q_{STG} lie under their respective limits. This region was investigated more carefully (see Fig. 53). The ranges of variation are entered in this diagram.

For this special fast diode the requirements are: Q_{STG} < < 30 µC, measured with a forward current i_F = 500 A and $-di_F/dt$ = 25 A/µsec, and U_F < 1.6 V for i_F = 500 A.

These requirements are satisfied for lifetimes 2.2 µsec < < τ_K < 3.1 µsec. The chosen lifetime must therefore lie within this tolerance range. The specification for τ_K would therefore be placed in the center of the range, i.e. at 2.6 µsec. The τ_K scattering of the gold diffusion procedure is small enough to keep the experimentally obtained values within the given limits.

Normally, the lifetimes obtained at a once chosen gold /129 diffusion temperature for various crystals from one and the same manufacturer (Wacker-Chemitronic) also lie within the permissible limits, and otherwise a slight temperature change (maximum ±10°C) is sufficient to bring the lifetimes into the prescribed range. In order to avoid yield losses, a specimen diffusion is generally undertaken, i.e. several slices from one crystal are gold-diffused and measured, from which any required temperature change can be determined.

In the manufacturing process given here for a fast power diode, the gold diffusion temperature is normally 880°C, and the diffusion time 1 hour. Changing the temperature by +5°C

shifted the mean value of τ_K to smaller values, but here too the distribution lay within the given limits.

The τ_K distribution obtained by this procedure proved to be sufficiently narrow for the application. Its width is substantially less than the width of the τ_K tolerance region 2.2 μsec $\leq \tau_K \leq$ 3.1 μsec.

The situation is less favorable for applying gold diffusion to frequency thyristors, because the τ_K tolerance region is smaller than that for diodes. Not only the forward voltage drop U_T and the recovery time t_q but also the ignition current i_{GT} of the gold diffusion are influenced. There are prescribed limits for all three parameters which must not be exceeded. In this situation, a specimen diffusion is required in every case, even more than one under certain circumstances, in order to place the mean value of τ_K in the narrow tolerance range.

7.7. Summary

The influences of gold diffusion on carrier lifetime and associated quantities are investigated for silicon diodes. A relationship between diffusion temperature and lifetime was obtained which is sufficiently precisely reproducible for the application. Temperatures greater than 750°C are required to shorten the lifetime. Gold diffusion for 1 hour at 890°C results in a typical lifetime value τ_K = 2.5 μsec. The influence of diffusion time is slight compared with that of the temperature. /130

The lifetime obtained depends on the capacity of the gold source at the surface of the silicon slice. This was shown by interrupted diffusions, in which the essentially inexhaustible Au reservoir on the surface was removed after a short preliminary diffusion.

The type of gold application (deposition from the vapor or coating in aqueous solution) has no influence on the lifetime obtained. Coating in aqueous solution was preferred.

Lifetime control in the development of a fast power diode (continuous limiting current 250 A) was described.

8. Summary /131

Technological and physical problems in control of carrier lifetime in silicon were studied. The investigations concentrated on controlling the lifetime of the free charge carriers

in the wide n-type base regions of gallium-diffused thyristor and diode structures in power electronics.

The information obtained and the development results can be summarized as follows:

a) The essential technological procedure for lifetime reduction is the indiffusion of gold as a recombination center. The gold diffusion process proved to be structure-dependent. The observed scattering of lifetime values from one crystal to the next after gold diffusion was attributed to the different vacancy generation rates. Radiochemical measurements yielded the following relationship for the largely position-invariant gold concentration C_S in the n-base regions of the slices as a function of time: $C_S = \sqrt{k_o t}$, where the dependence of the constant k_o on the real structure of the crystal causes the scattering of the gold concentration and lifetime values. The scattering in the lifetime values can be reduced by a two-step diffusion process.

b) In the production of fast power diodes, the required reduction in lifetime can be carried out by indiffusion of gold in a fashion sufficiently reproducible for the application, when uniform crystal material (same producer, same procedure) is used. The type of gold application has no influence on the resulting lifetime. Coating the slices in aqueous solution (chloroauric acid) under the action of ultrasonics proves to be a very logical procedure for applying the gold. Temperatures above 750°C are required to shorten the lifetime. The influence of diffusion time is slight in comparison with that of temperature. The lifetime is controlled by the choice of diffusion temperature, using a constant diffusion time of 1 hour. One-hour gold diffusion at 890°C results in a typical lifetime value of 2.5 µsec. /132

c) A practically feasible technological procedure to increase lifetime is the phosphorus gettering process. It finds application to components (e.g. thyristors with high reverse voltage rating), for which the lifetime after gallium diffusion is too small. Even when the gettering process is carried out at relatively low temperatures below 1000°C, it works very well. Traces of heavy metals which enter during the diffusion treatment are removed from the interior of the Si slices. The gettering mechanism consists in the formation of heavy-metal-phosphorus compounds (e.g. Au_2P_3) in the n^+-type region produced at the surface through phosphorus diffusion. The effectiveness of the gettering procedure and the remaining traces of heavy-metal impurities are checked radiochemically. The lifetimes attainable by phosphorus gettering are limited by the appearance of other impurities not belonging to the heavy metals and which cannot be gettered with phosphorus.

d) Lifetime values up to 300 μsec are achieved in diffused slices by means of a new gallium diffusion procedure which exploits the gettering and doping properties of liquid gallium layers on silicon. This procedure combines doping and gettering into a single step. The procedure by which the slices are cleaned before diffusion, the purity of the gallium employed, and the cleanness of the vicinity of the slices during diffusion all have an astonishingly small influence on the resulting high lifetimes.

e) The temperature variation of carrier density in p-type and n-type conducting silicon slices (10-10,000 Ωcm) was determined from measurements of the Hall effect in the temperature range between 100 K and 480 K. Analysis of the concentration-temperature curve makes it possible to determine the concentration of deep impurities even in cases in which radiochemical methods fail. /133

In ordinary n-type conducting zone-pulled starting crystals for component production, there are already deep impurities in varying concentrations (some 10^{12} cm^{-3}). This initial level of deep impurities has dropped in recent years, as studies in crystals of various "ages" have shown.

Gallium diffusion (Si slices about 70 hours in Ga vapor in an argon atmosphere at about 1200°C) is always associated with penetration of relatively rapidly diffusing deep impurities. The concentrations of these deep impurities (several 10^{13} cm^{-3}) far surpass those present in the starting material. The injected impurities are donors of a type with an activation energy of $E_{CD} = 0.28 \pm 0.01$ eV, as shown by analysis of the log n vs. 1000/T diagram. The deep donors are homogeneously distributed in the base regions but occur in different concentrations from one slice to the next. They affect the carrier lifetime. The deep donor cannot be detected radiochemically. It is probably an element of the second or sixth group of the periodic table. The strongest suspicion falls on sulfur. The source of the impurity is sought outside the preparation of the slices and the ampules. The results support the hypothesis that the permeability of the quartz ampule plays a large role at the high diffusion temperatures.

The level of deep impurities in the component is decisively determined by the gallium diffusion (gallium vapor). The further steps in thyristor production, such as oxidation in moist oxygen and the subsequent phosphorus diffusion, are not accompanied by any substantial changes in the impurity level.

Gettering produces marked decreases in the impurity level. /134
However, only some of the incorporated deep impurities can be eliminated.

Highly resistive n-type and p-type conducting slices, subjected to the production process together with normal n-type silicon slices, are well suited to check the impurity level during the production process. Unknown penetrating impurities can be sensitively detected by changing the net doping in these probe slices.

f) The decay time method can be used to measure the lifetime in the n-base region of p^+np^+ and $n^+p^+np^+$ slices with comparatively little preparative effort. In particular, the method is applicable to finished thyristors. The measurement proceeds as follows: An initially constant and then exponentially decaying voltage is applied to the slices to be tested, which must be held at elevated temperature. The current through the slice is observed. Under certain conditions, it passes through a minimum. The time interval between this minimum and the beginning of the voltage decay is the decay time t_d.

The new lifetime measuring method was checked on a large number of p^+np^+ and $n^+p^+np^+$ slices. As a comparison, the lifetime was measured for the same specimens by the retention time method. The measured value of t_d depends on, among other things, the resistivity of the n-base region, the temperature, and the time constant of the voltage decay.

This procedure can be applied in production in quantity of slices of a specific type. With appropriate choice of measuring conditions, there is an approximately linear relationship between t_d and the lifetime. Separate calibration is required for each type of slice.

The method is not applicable to slices with short lifetimes achieved by gold diffusion. Such slices behave completely dif- /135
ferently in the decay time measurement from slices with short lifetimes without gold diffusion. The differences in behavior between gold-diffused and non-gold-diffused slices with the same short lifetimes is interpreted to mean that a recombination center other than gold predominates in non-gold-diffused slices.

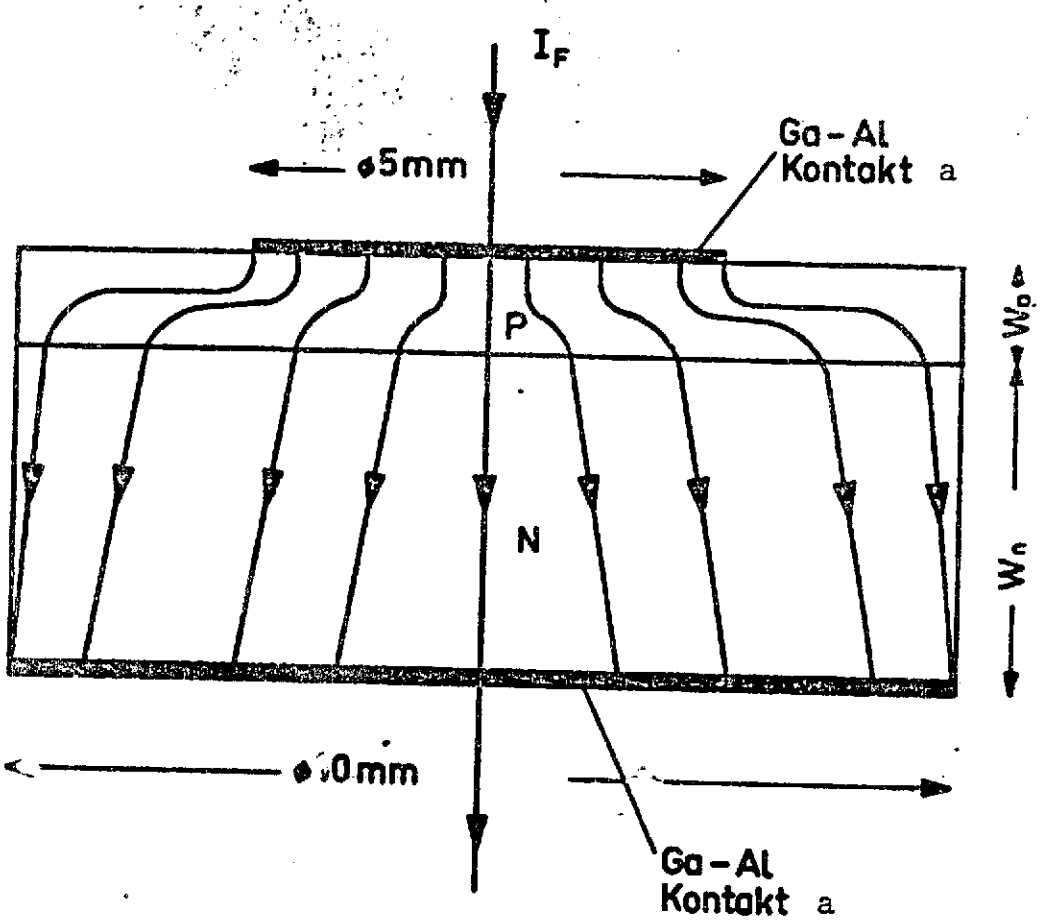
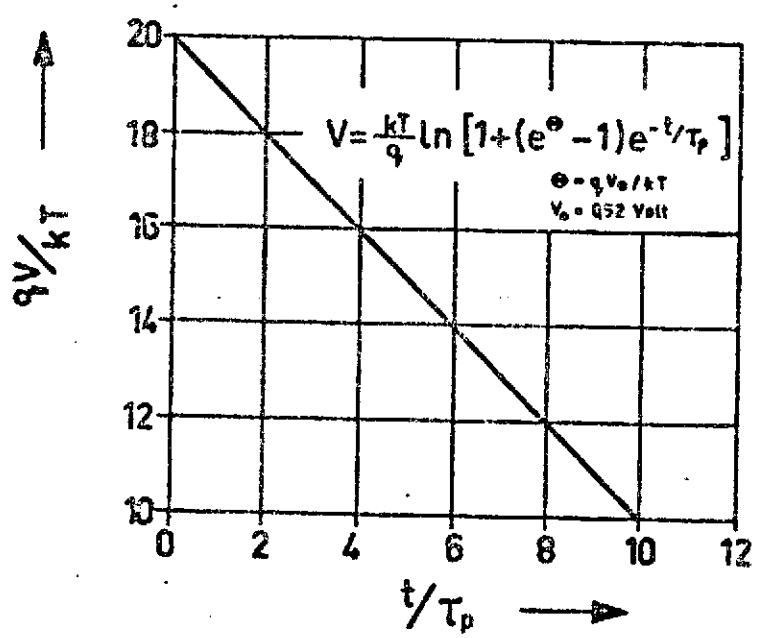


Fig. 1. Schematic representation of diodes employed.

Key: a. Contact
ø = diameter

(a)



/143

Fig. 2(a). Post-injection voltage V as a function of time under simple conditions.

(b)

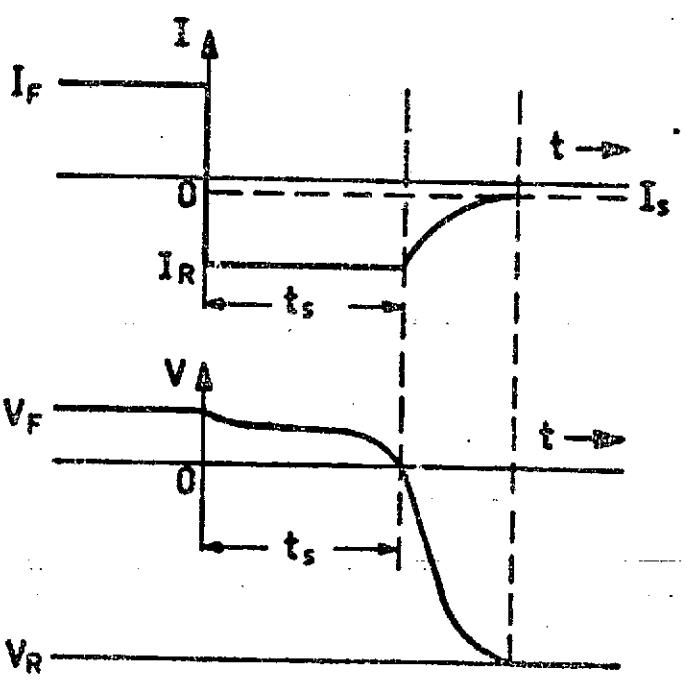
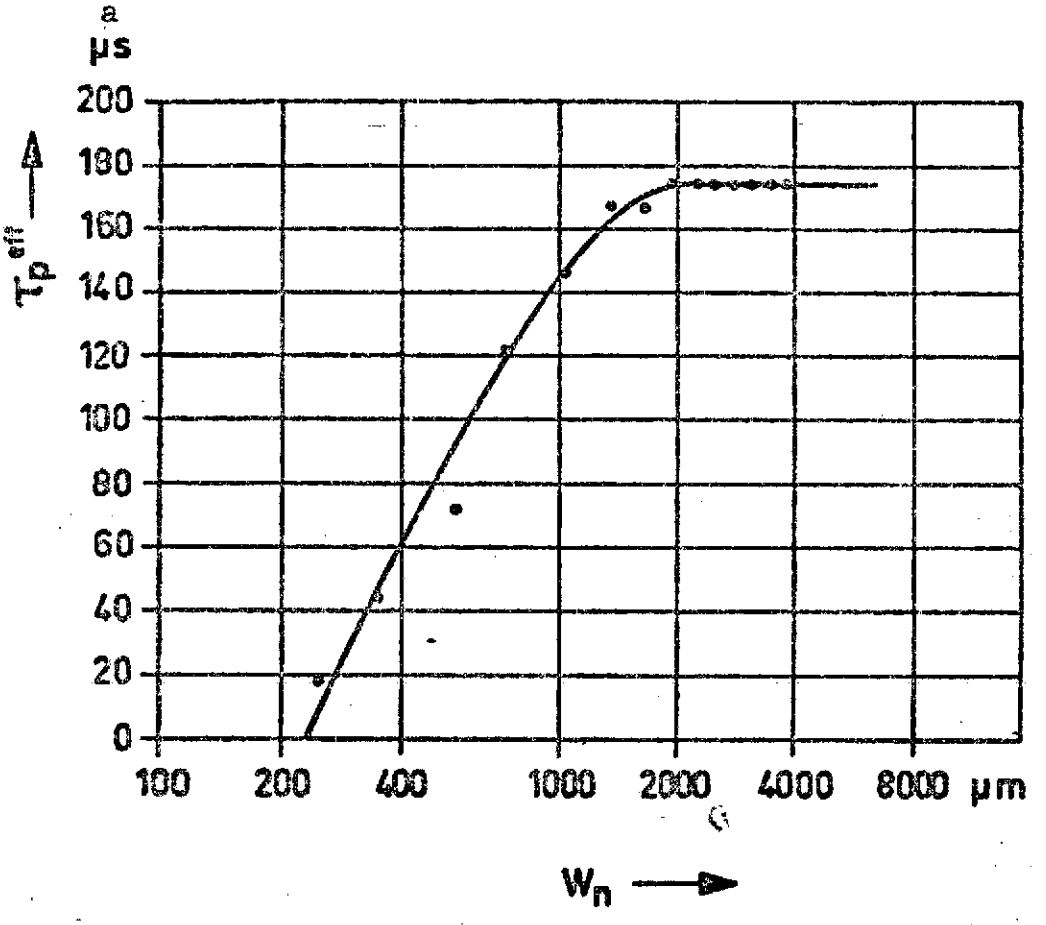


Fig. 2(b). Current and voltage curves in the retention time method.



$$\tau_p^{eff} = 174\ \mu s^{a}\ bei^{b}\ W_n > 2000\ \mu m$$
$$= 4.35 L_p^{eff}$$
$$L_p^{eff} = 460\ \mu m$$

Fig. 3. Measured relationship between τ_p^{eff} and thickness W_n of diode base.

Key: a. µsec
b. For

/145

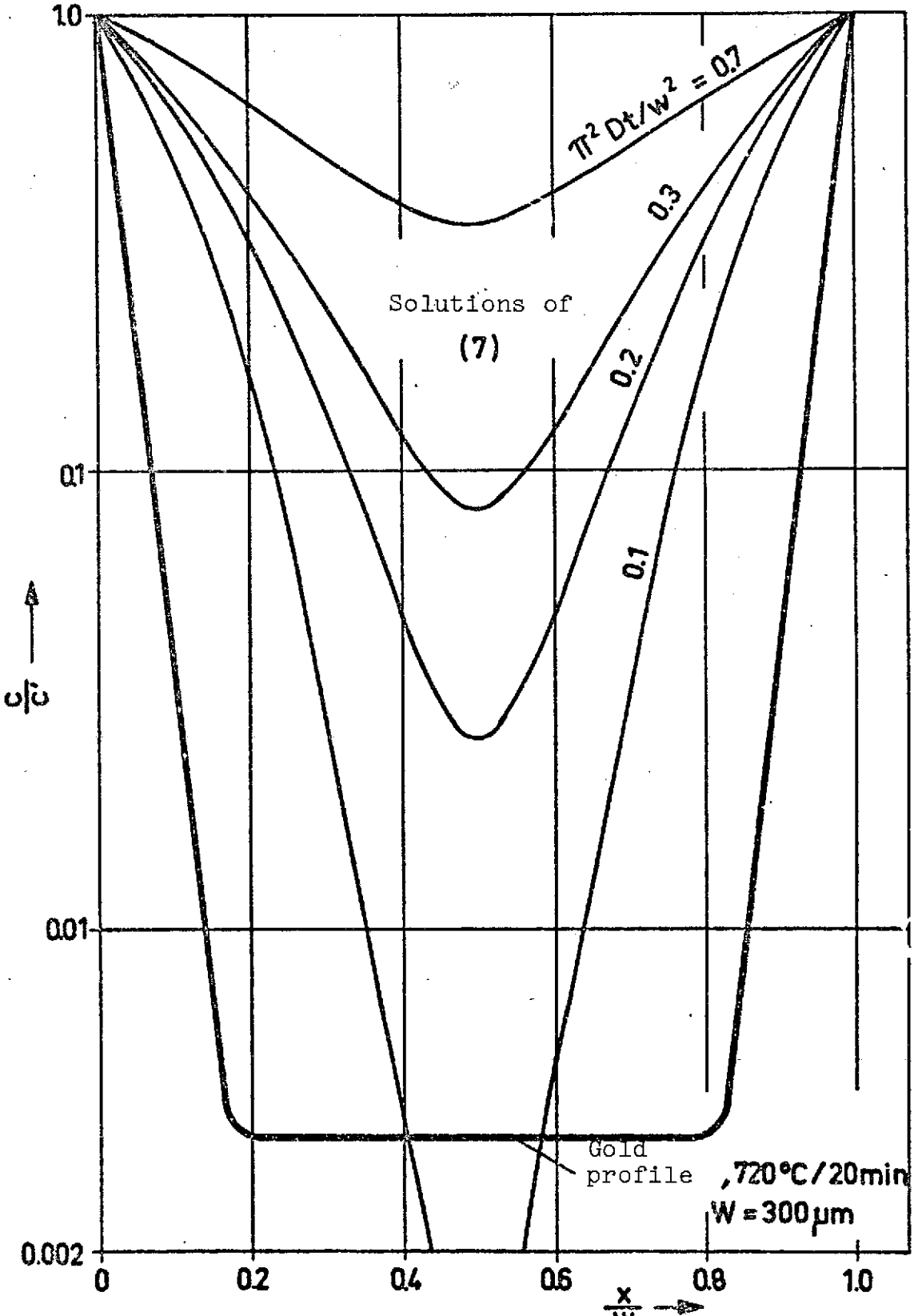


Fig. 4. Comparison between ideal diffusion profiles derived from (7) and experimentally determined gold diffusion profile.

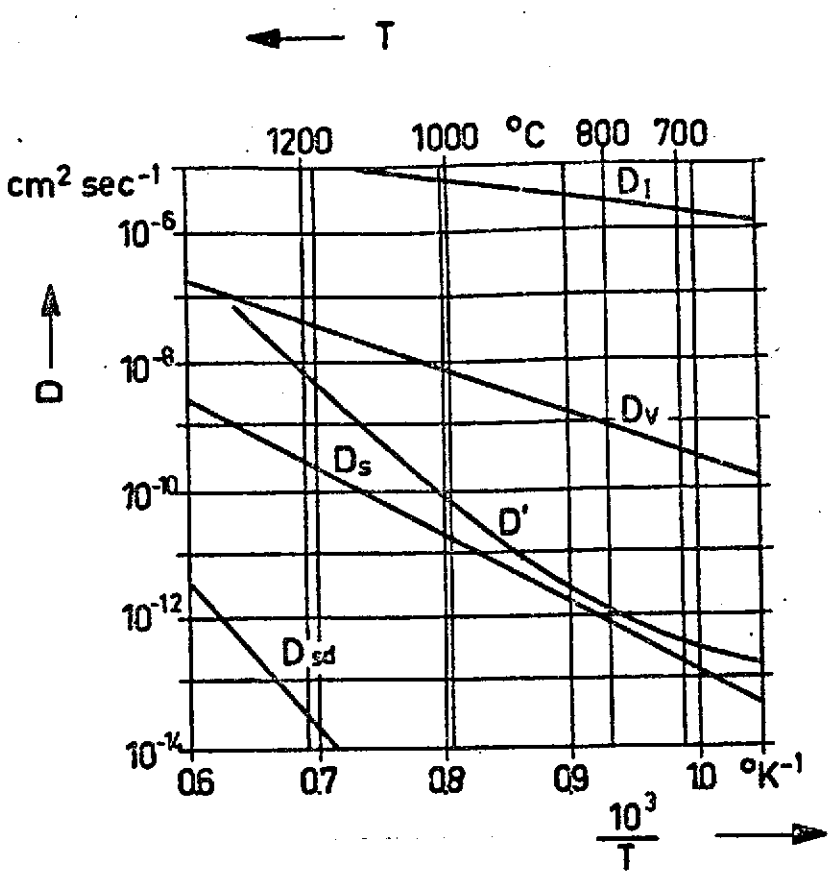


Fig. 5. Diffusion of gold in silicon.

D_{sd} = self-diffusion constant as defined by [23]
D_V = double vacancy diffusion constant
D_I = diffusion constant for gold in interlattice diffusion
D_S = constant for diffusion of gold over lattice sites
D' see (17)

/147

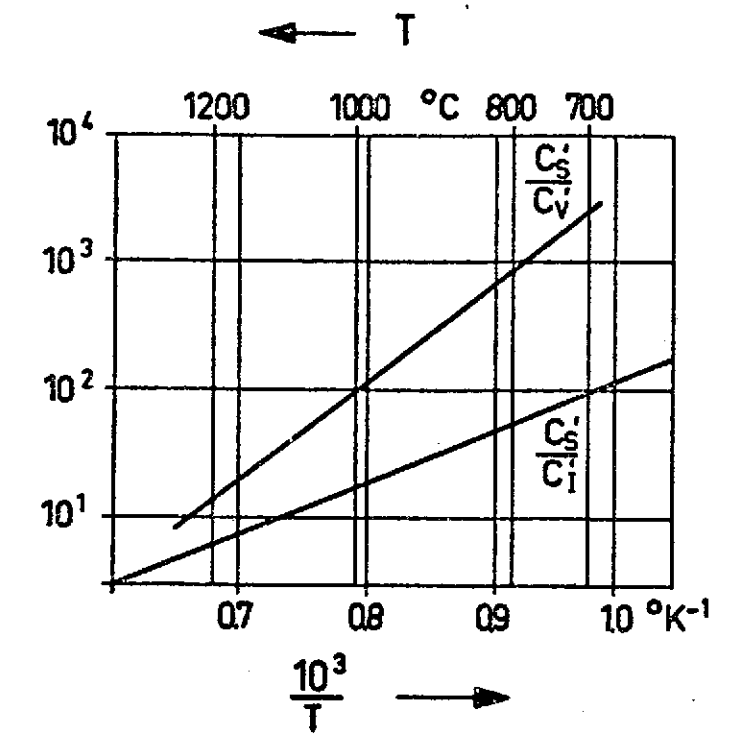


Fig. 6. C_S'/C_V' and C_S'/C_I' as functions of T [22, 23, 24].

/148

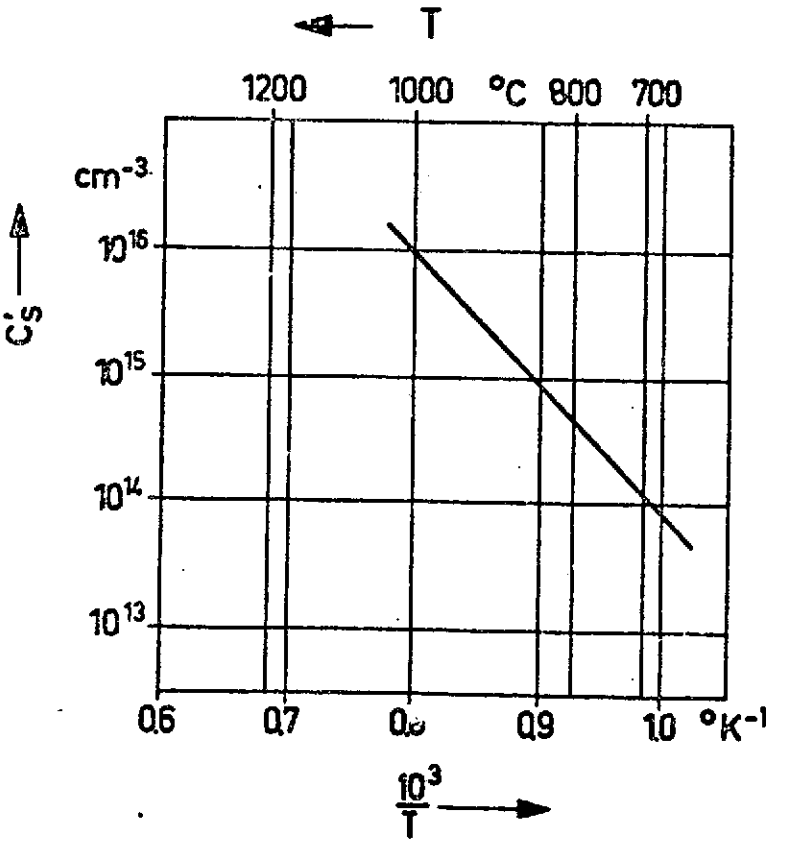


Fig. 7. C_S' as function of T [24].

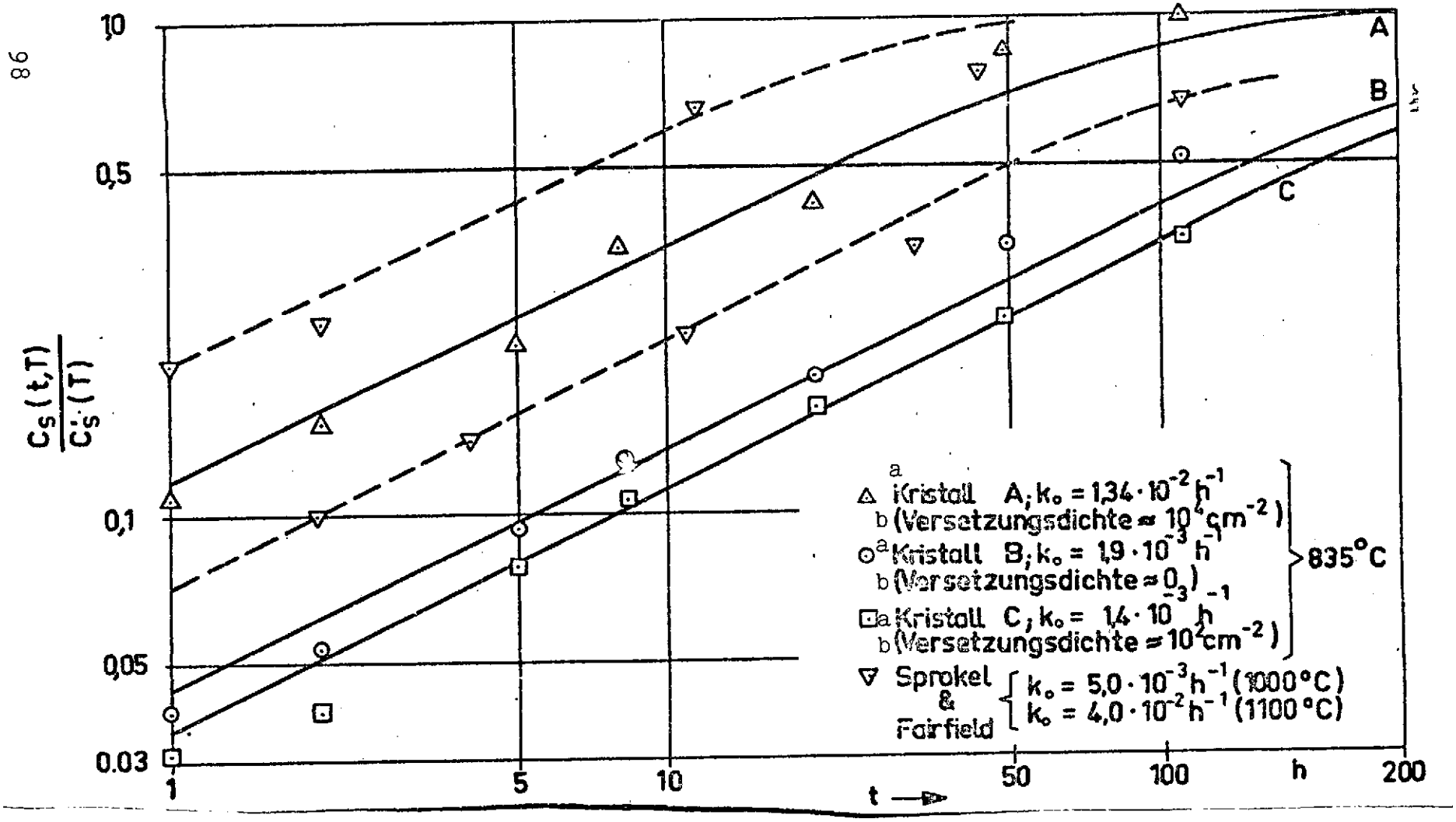


Fig. 8. Gold concentration level in center of specimen as function of time for different crystals. Broken and solid curves given by $\frac{C_s(t,T)}{C_s'(T)} = \sqrt{\left(1 - e^{-k_o t}\right)}$

Crystals A, B: zone-pulled
Crystal C: crucible-pulled

Key: a. Crystal; b. dislocation density; h = hour.

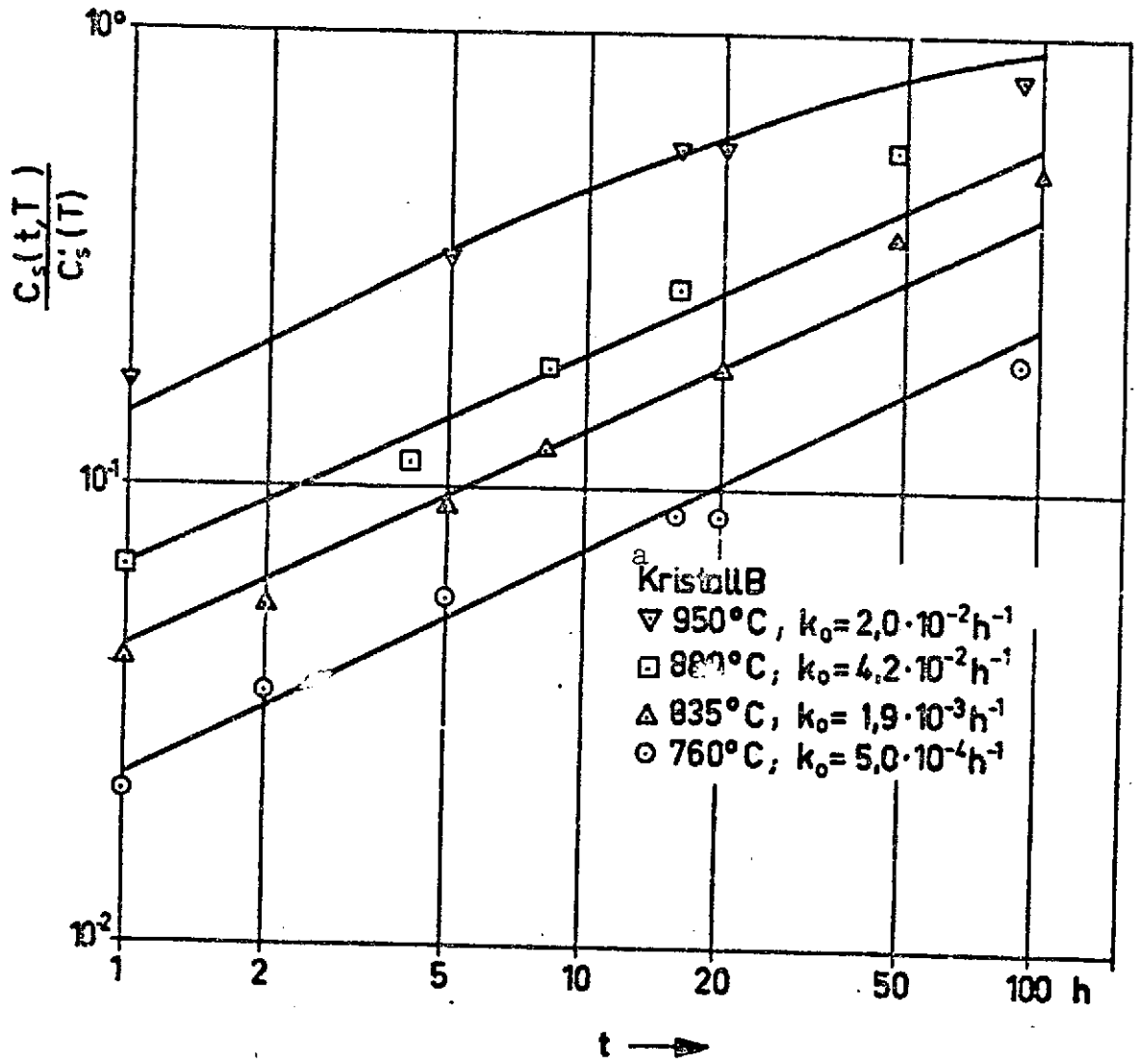


Fig. 9. Gold concentration level in center of specimen as function of time for different diffusion temperatures.

Key: a. Crystal
h = hour

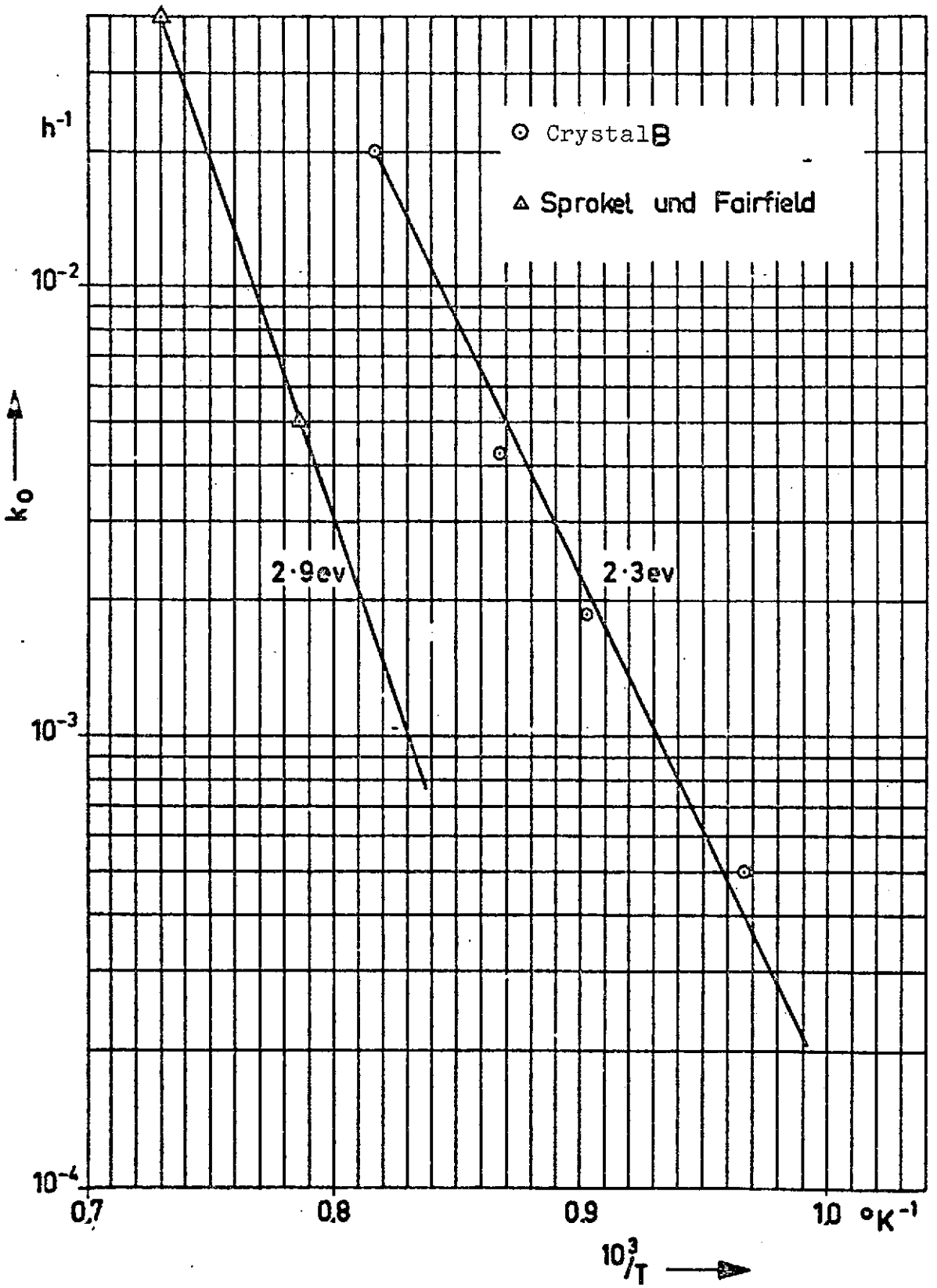


Fig. 10 Parameter k_o as function of T.

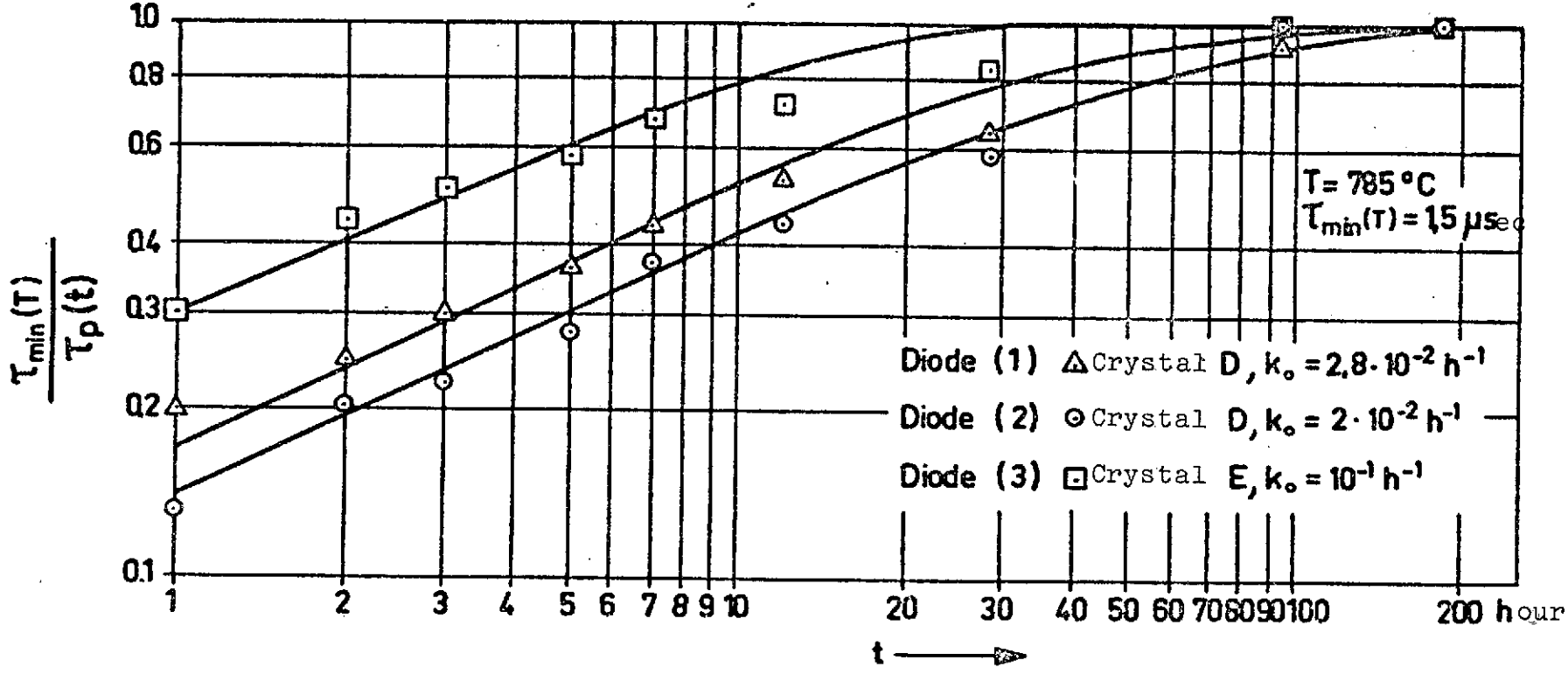


Fig. 11. τ_p as a function of the duration t of gold diffusion.
Solid curve given by $\frac{\tau_{min}}{\tau_p(t)} = \sqrt{(1 - e^{-k_o t})}$

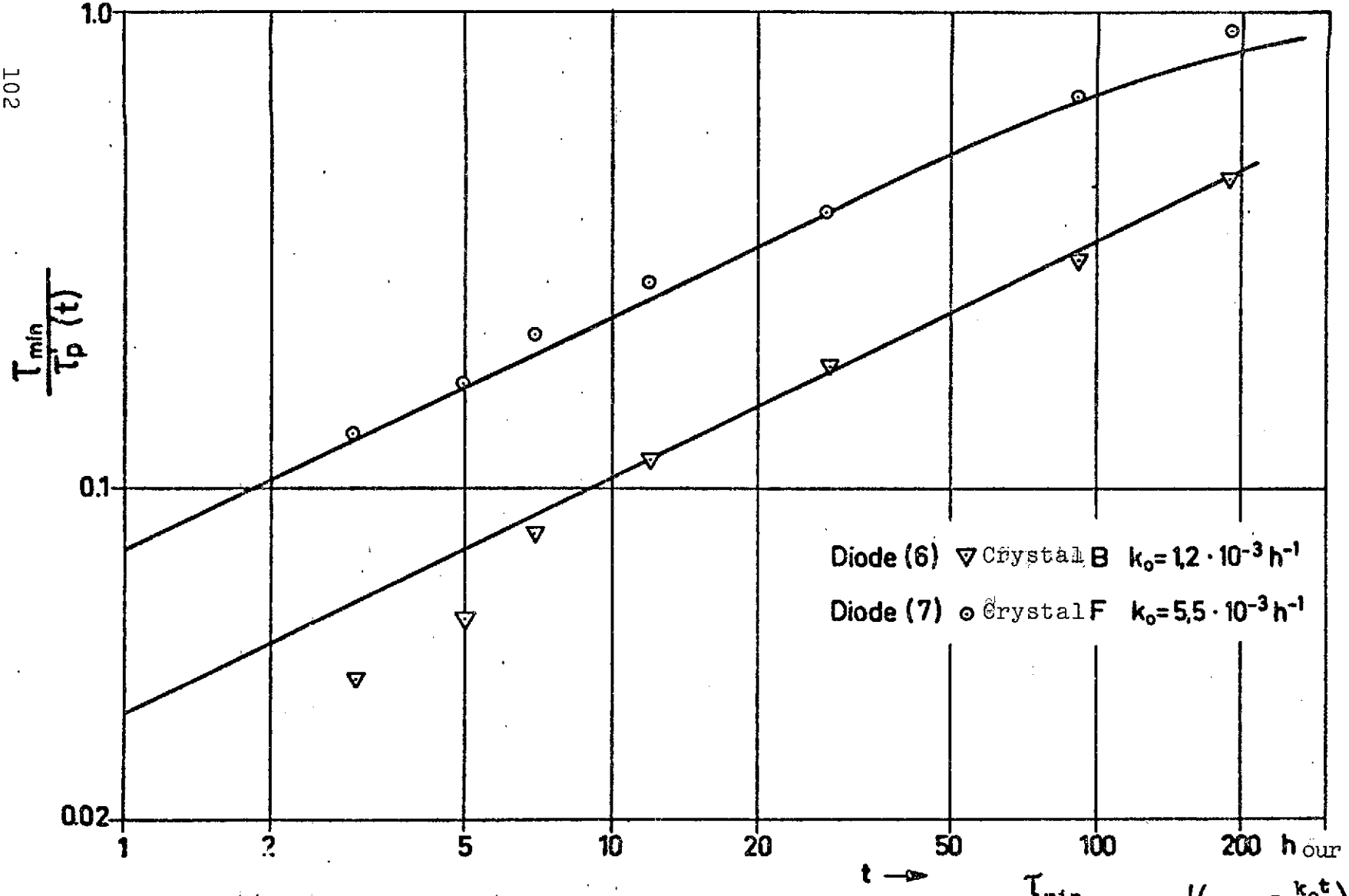


Fig. 12. τ'_p as a function of t, from (29). Solid curve given by $\frac{\tau_{min}}{\tau'_p(t)} = \sqrt{\left(1 - e^{-k_o t}\right)}$

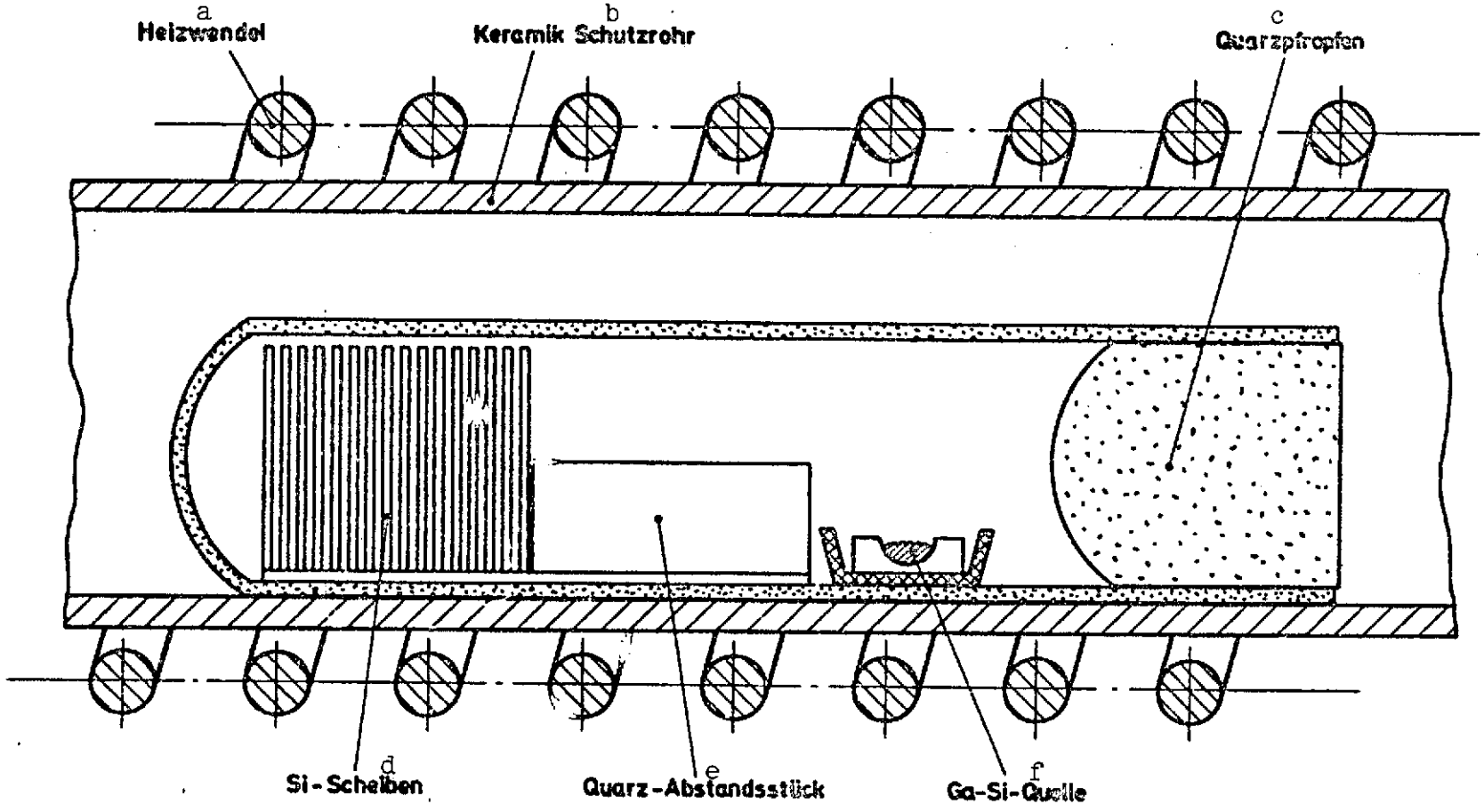


Fig. 13. Gallium diffusion ampule.

Key: a. Heating coil
b. Ceramic protective pipe
c. Quartz stopper
d. Si slices
e. Quartz spacer
f. Ga-Si source

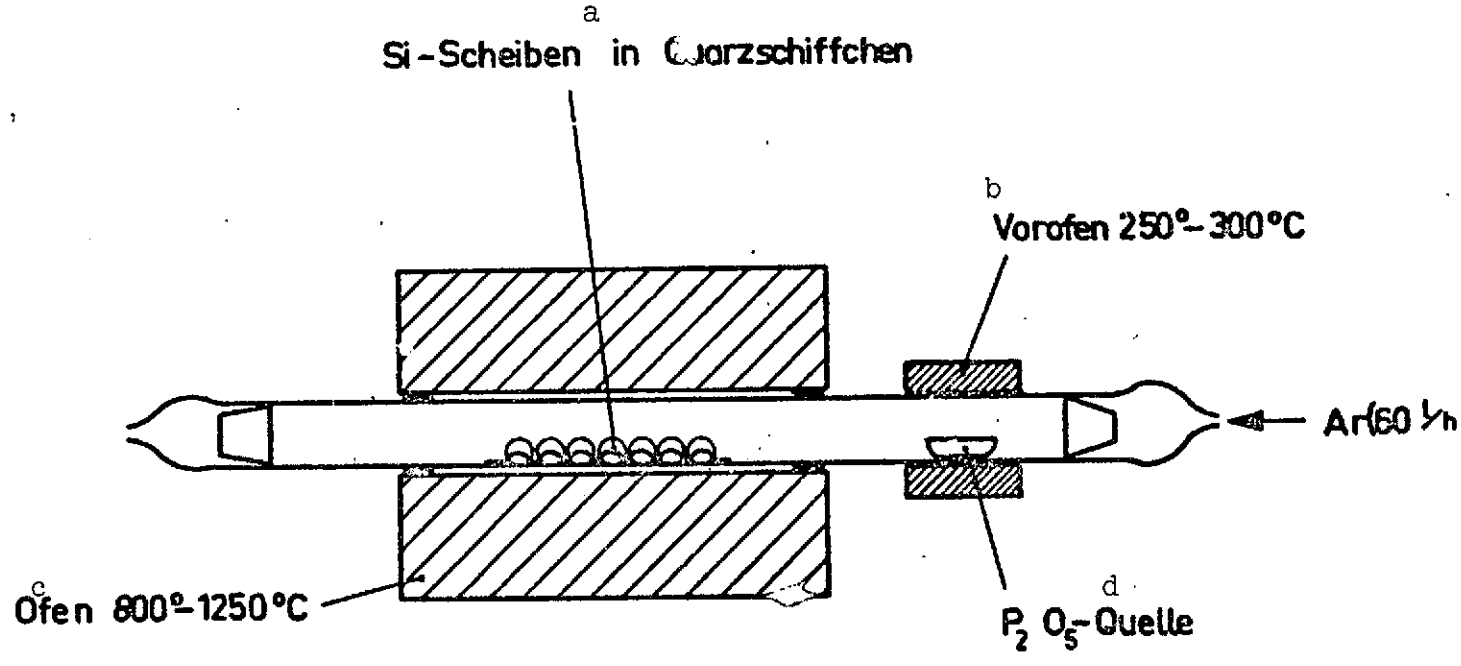


Fig. 14. Depiction of P_2O_5 gettering.

Key: a. Si slices in quartz boat
b. Preheating furnace
c. Furnace
d. Source
l/h = liter/hour

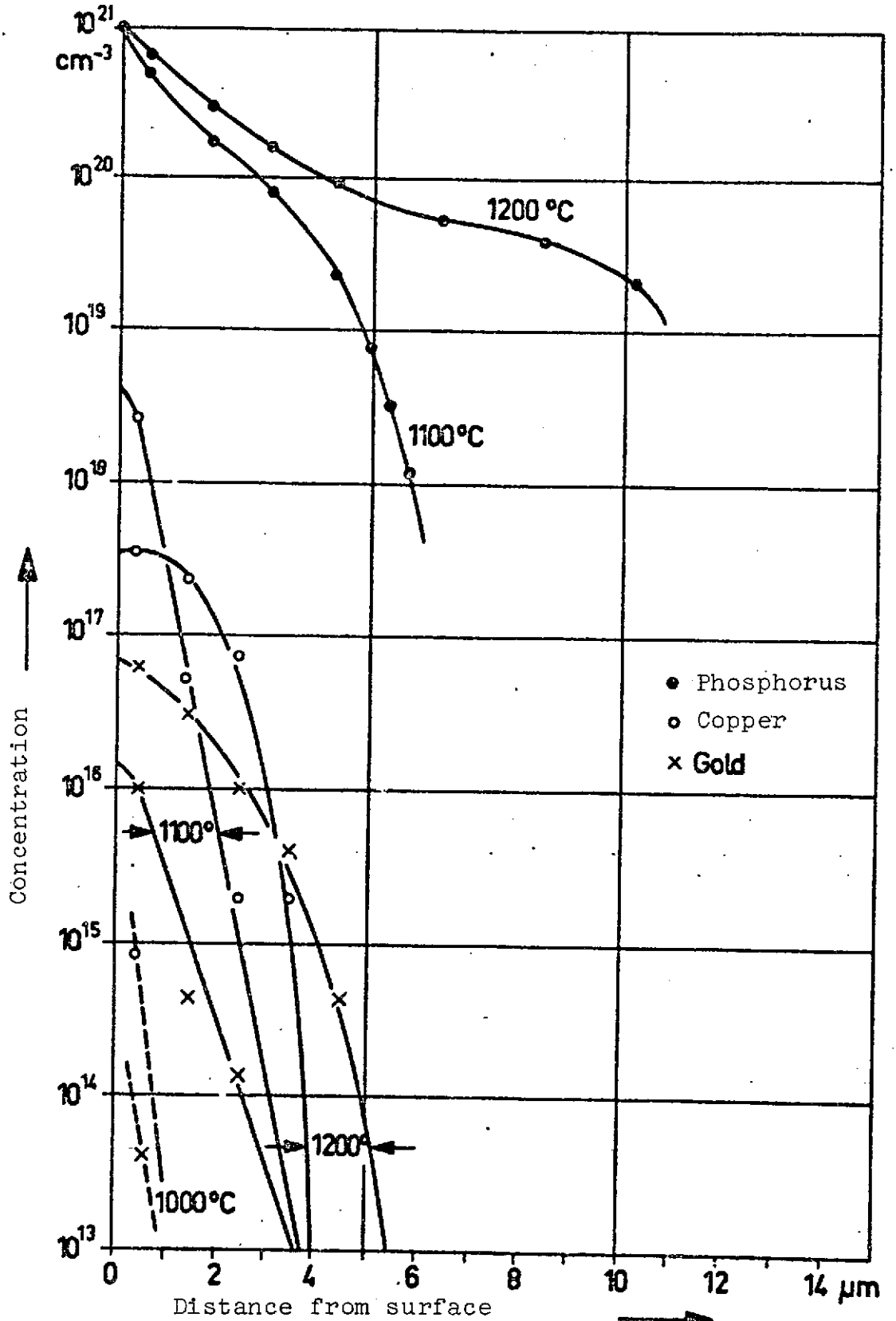


Fig. 15. Gold and copper profiles with associated phosphorus profiles.

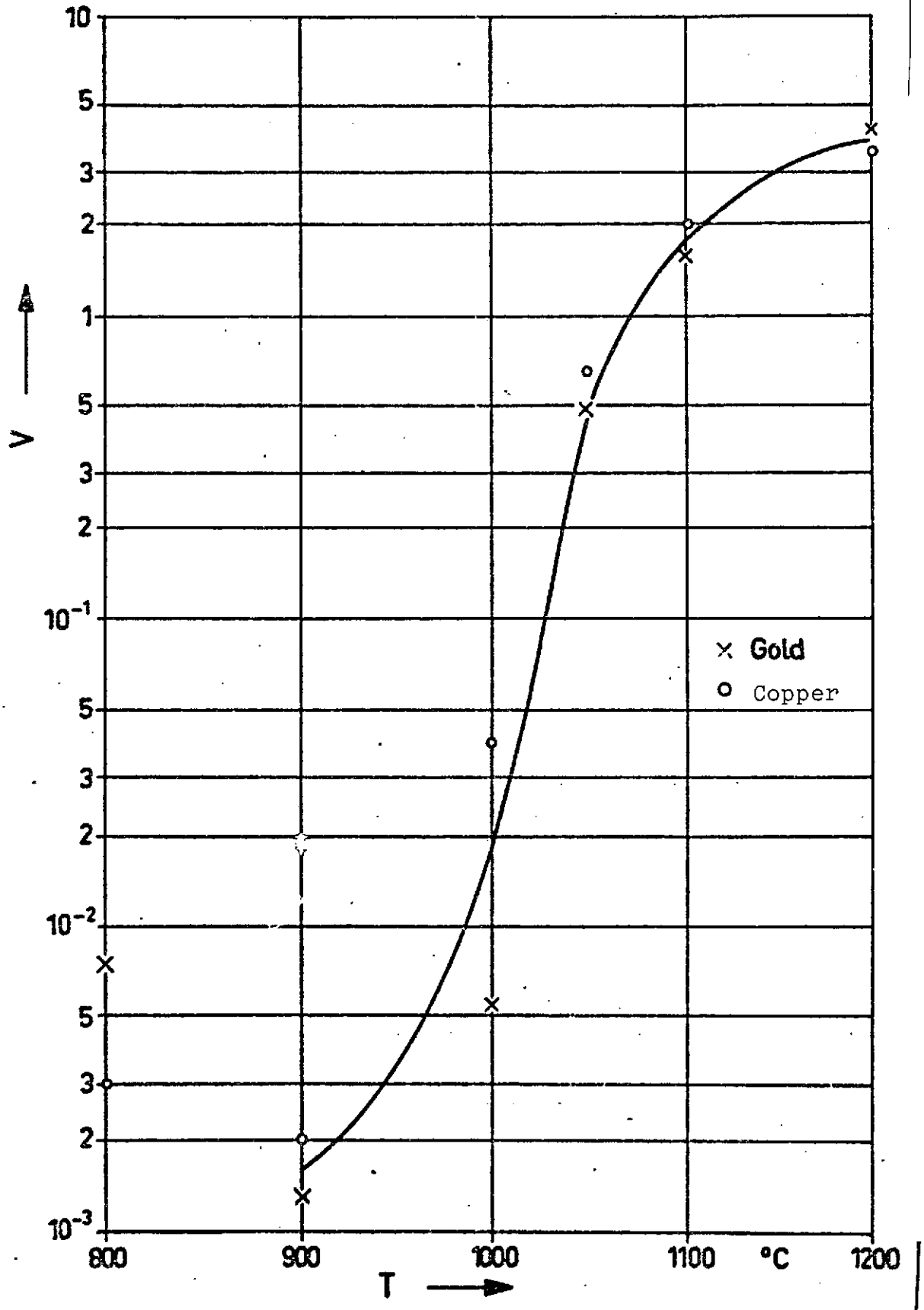
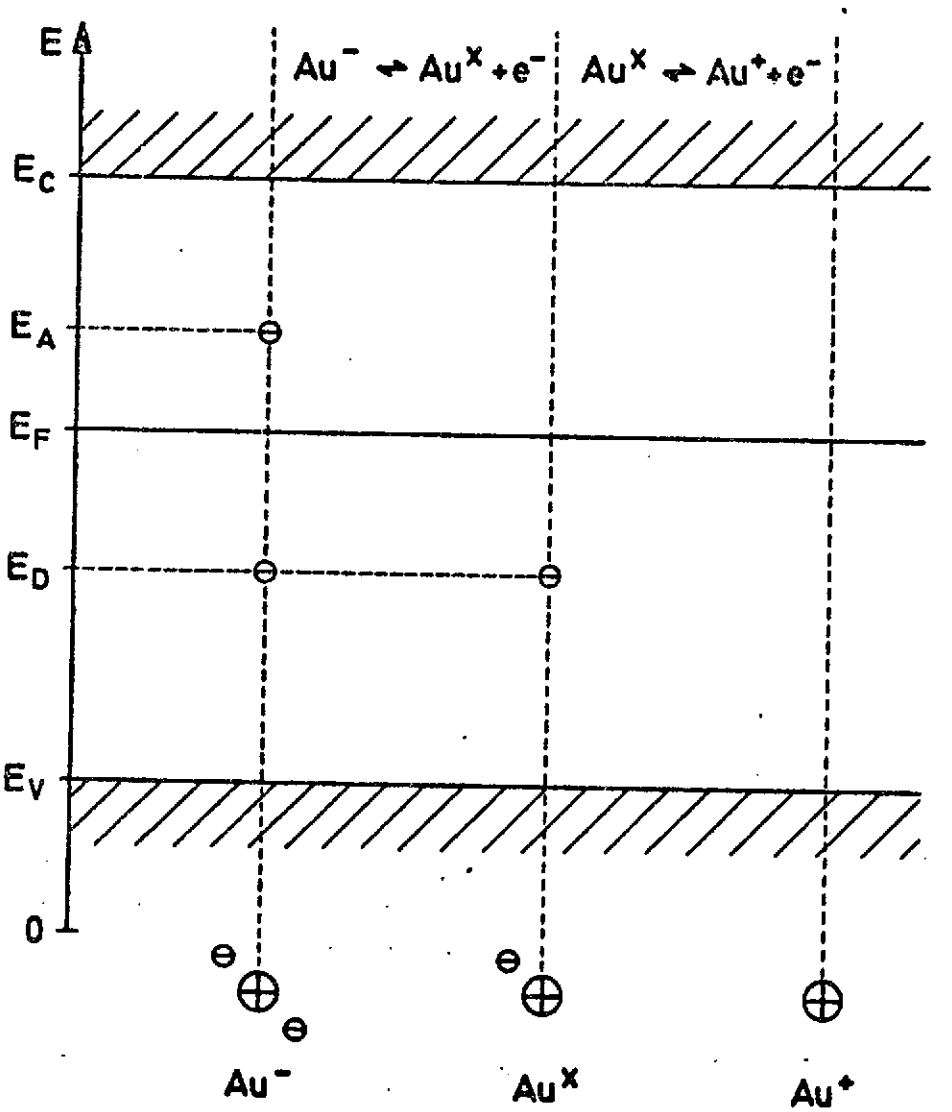


Fig. 16. Distribution V of metallic impurities as a function of gettering temperature.

/158



$E_C - E_V = 1.13 \pm 0.04\,eV$

$E_C - E_A = 0.54 \pm 0.02\,eV$

$E_D - E_V = 0.35 \pm 0.02\,eV$

Fig. 17. Gold levels in silicon.

/159

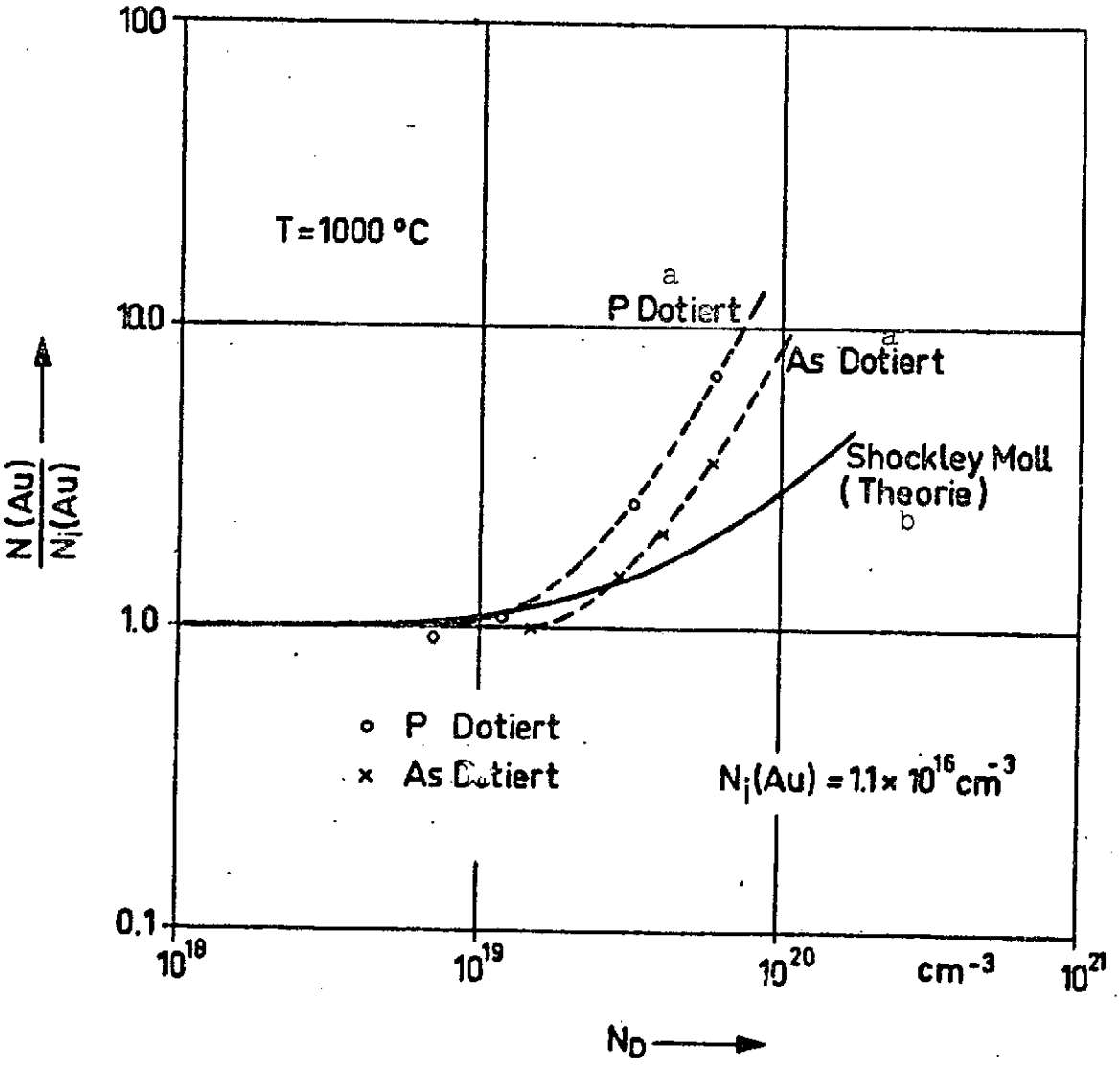


Fig. 18. Solubility of gold as function of donor concentration [43].

Key: a. Doped
b. Theory

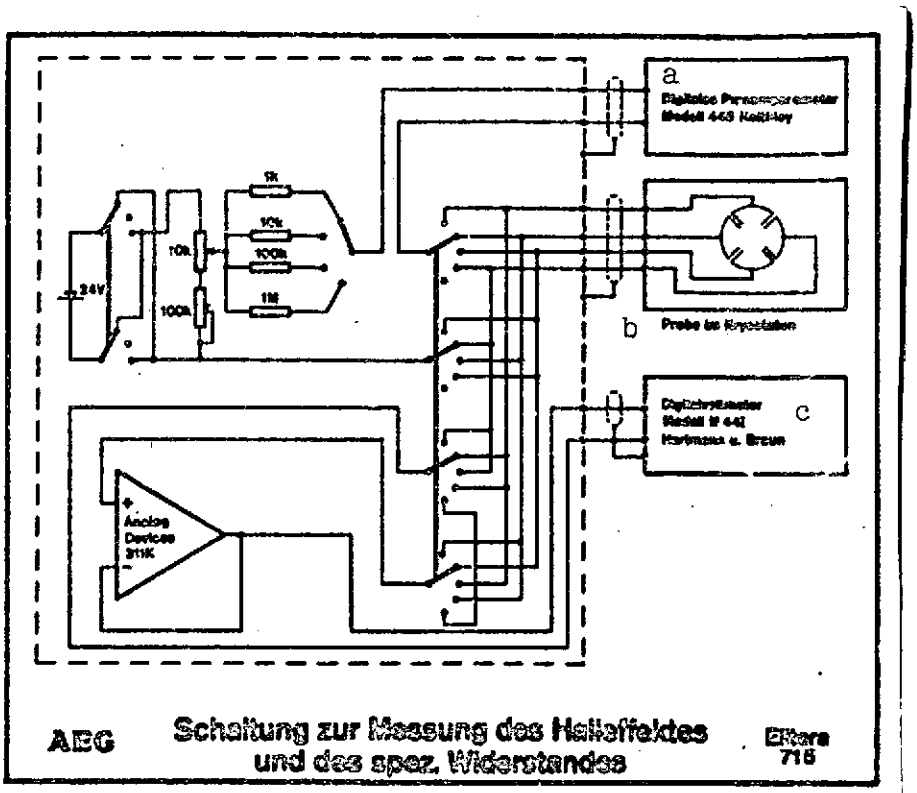


Fig. 19. Circuit for measuring Hall effect and resistivity.

Key: a. Digital picoammeter model 446 Keithley
b. Specimen in cryostat
c. Digital voltmeter model 4-4I Hartmann and Braun

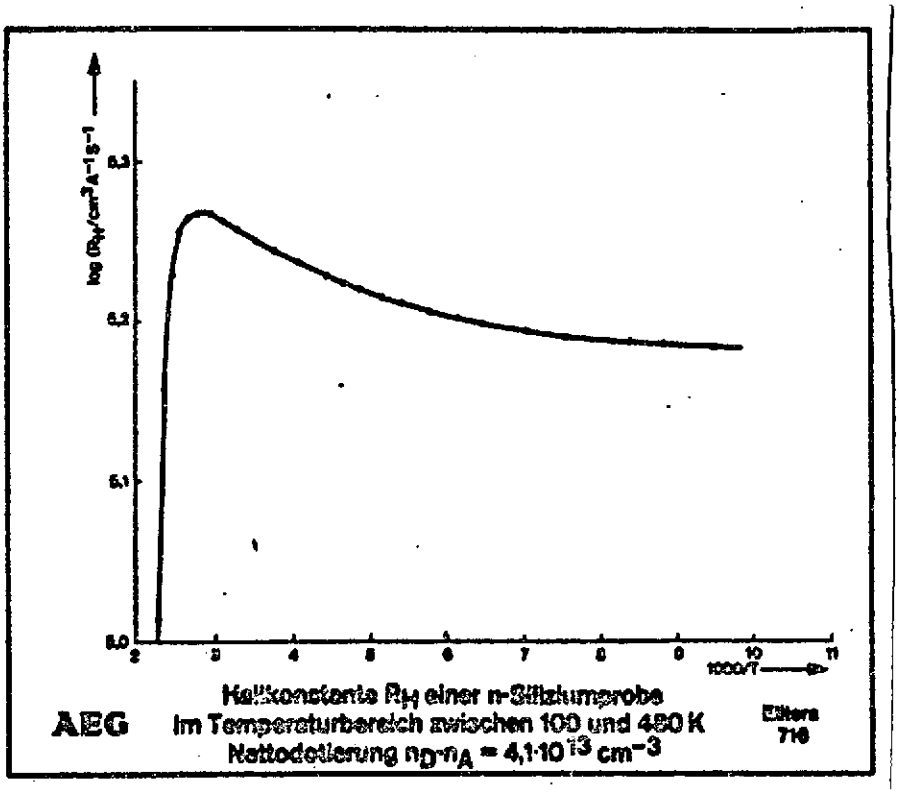


Fig. 20. Hall constant R_H of an n-type silicon specimen in the temperature range between 100 and 480 K, net doping $n_D \cdot n_A = 4.1 \cdot 10^{13}$ cm^{-3}.

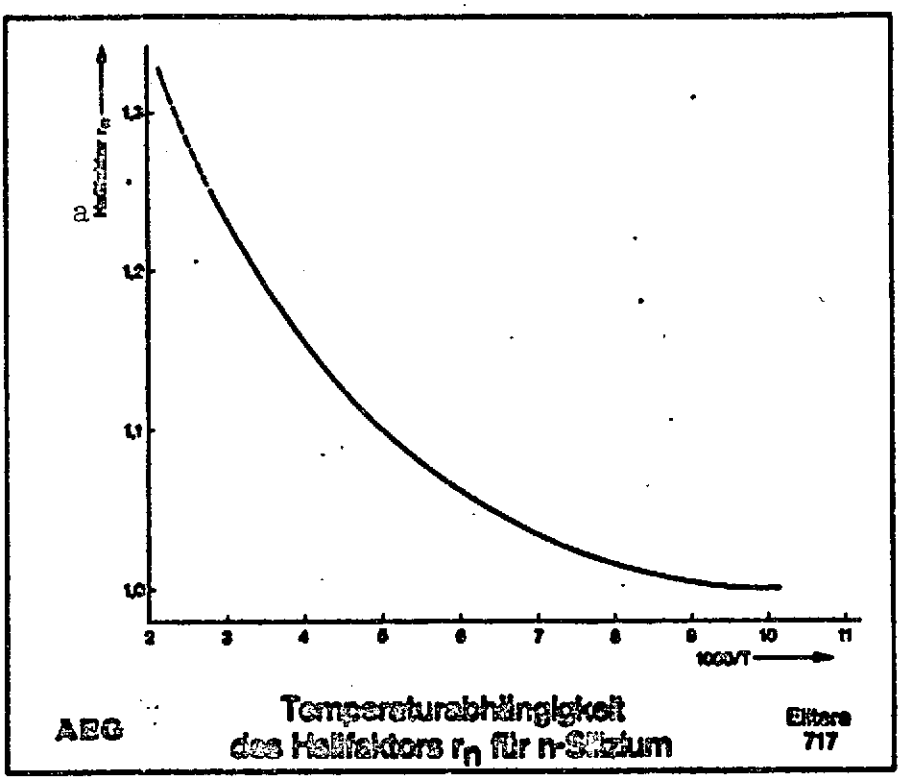


Fig. 21. Hall coefficient r_n for n-type silicon as function of temperature.

Key: a. Hall coefficient

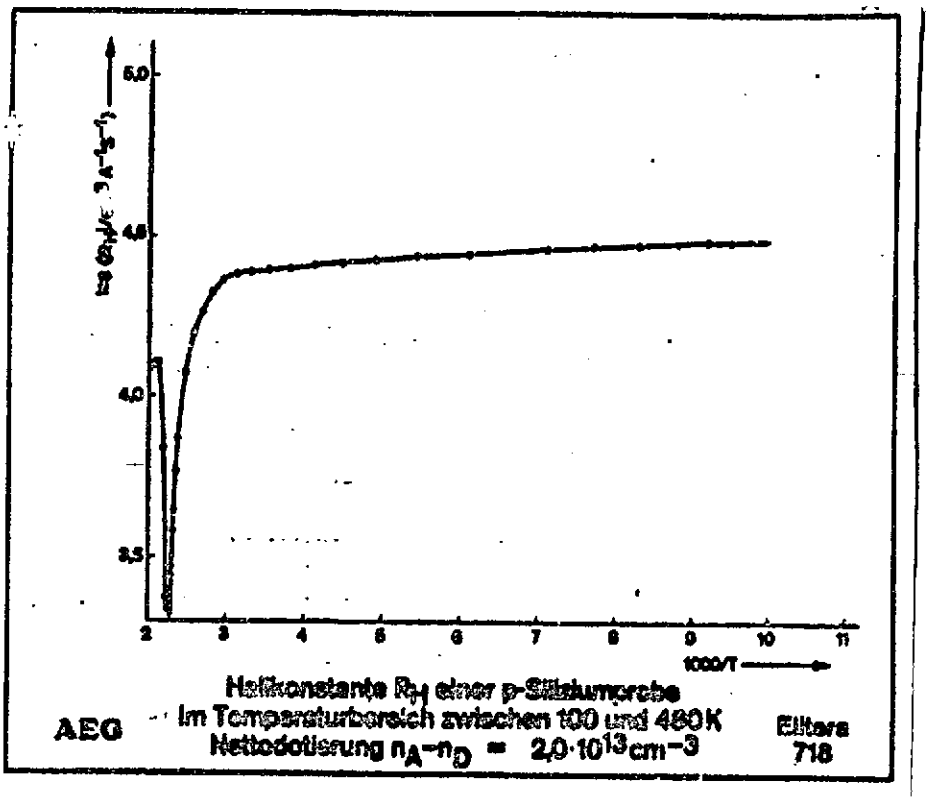


Fig. 22. Hall coefficient R_H of a p-type silicon specimen in temperature range between 100 and 480 K. Net doping $n_A - n_D$ = $= 2.0 \cdot 10^{13}$ cm^{-3}.

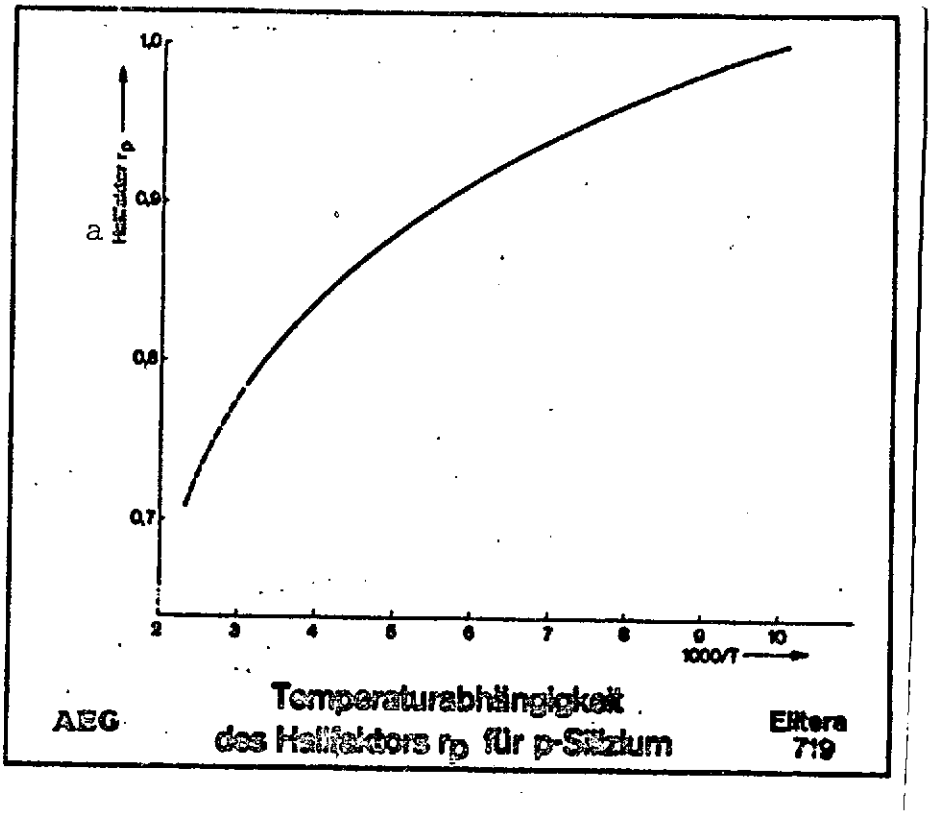


Fig. 23. Hall factor r_p for p-type silicon as function of temperature.

Key: a. Hall coefficient

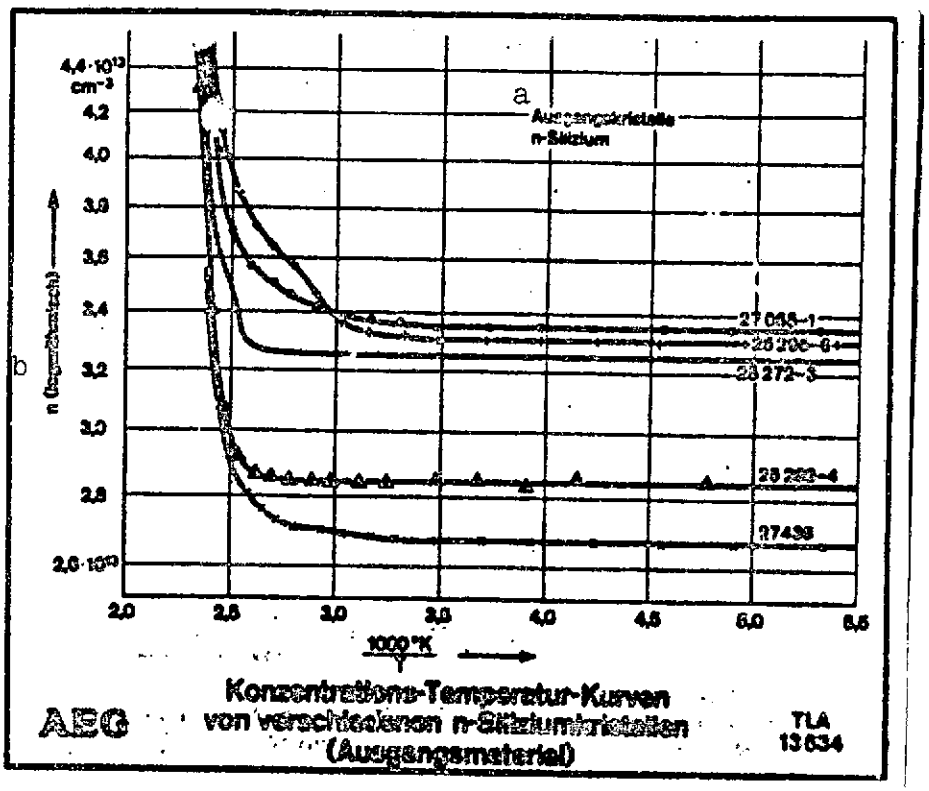


Fig. 24. Concentration-temperature curves of various n-type silicon crystals (starting material).

Key: a. Starting crystals n-type silicon
b. Logarithmic

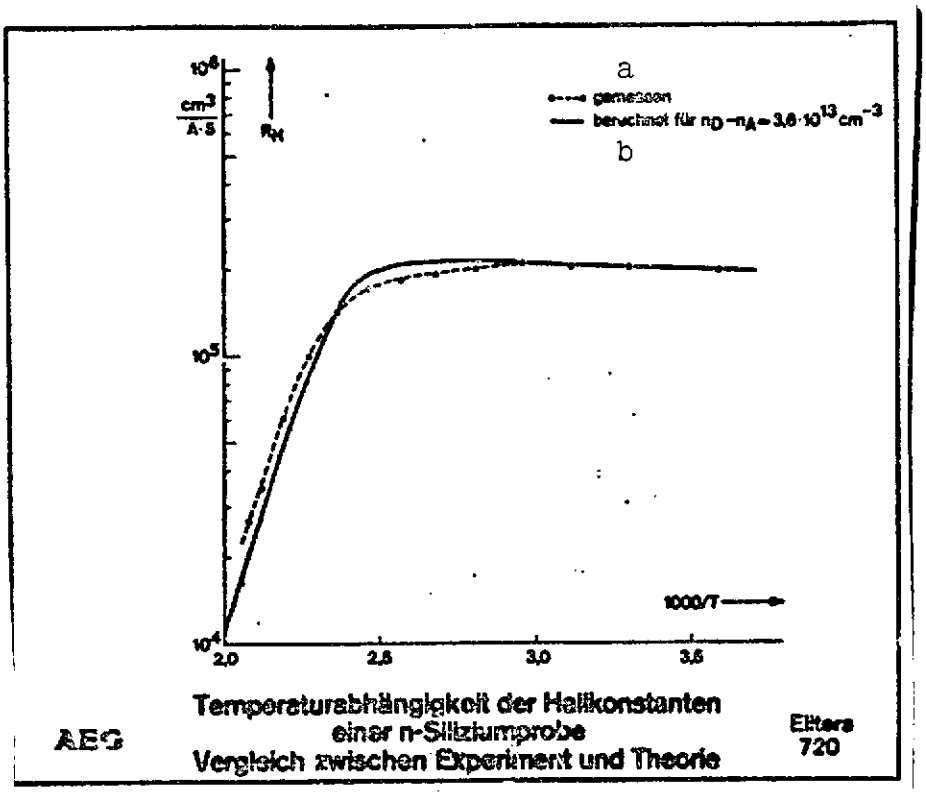


Fig. 25. Hall constant as function of temperature for an n-type silicon specimen. Comparison between experiment and theory.

Key: a. Measured
b. Calculated for
s = sec

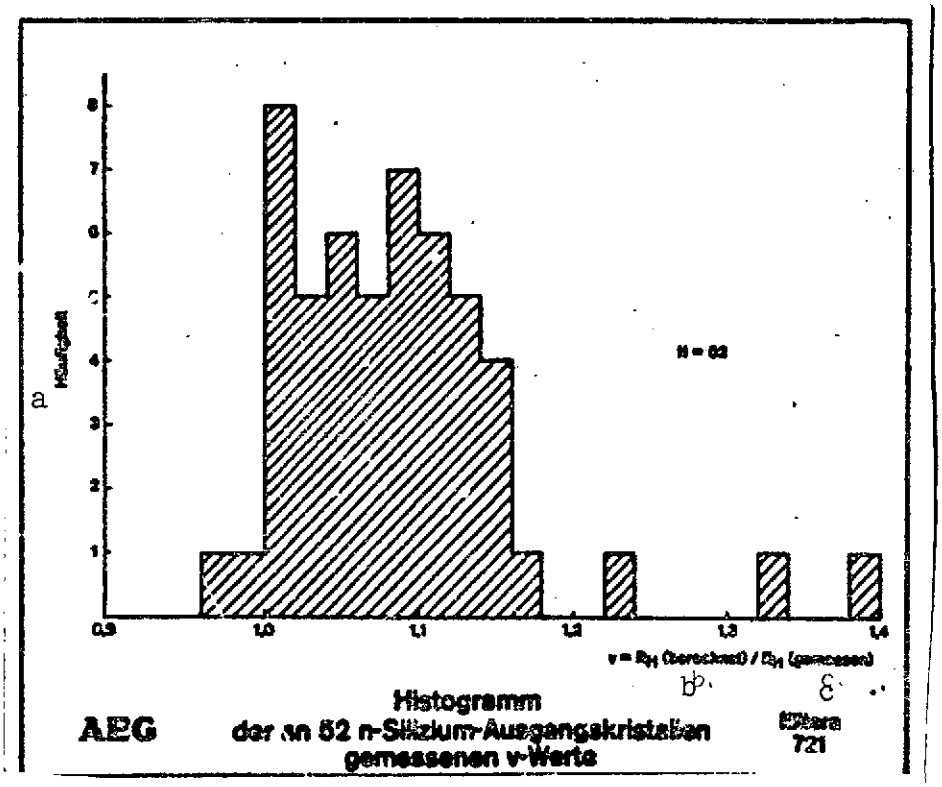


Fig. 26. Histogram of v-values measured on 52 n-type silicon starting crystals.

Key: a. Frequency
b. Calculated
c. Measured

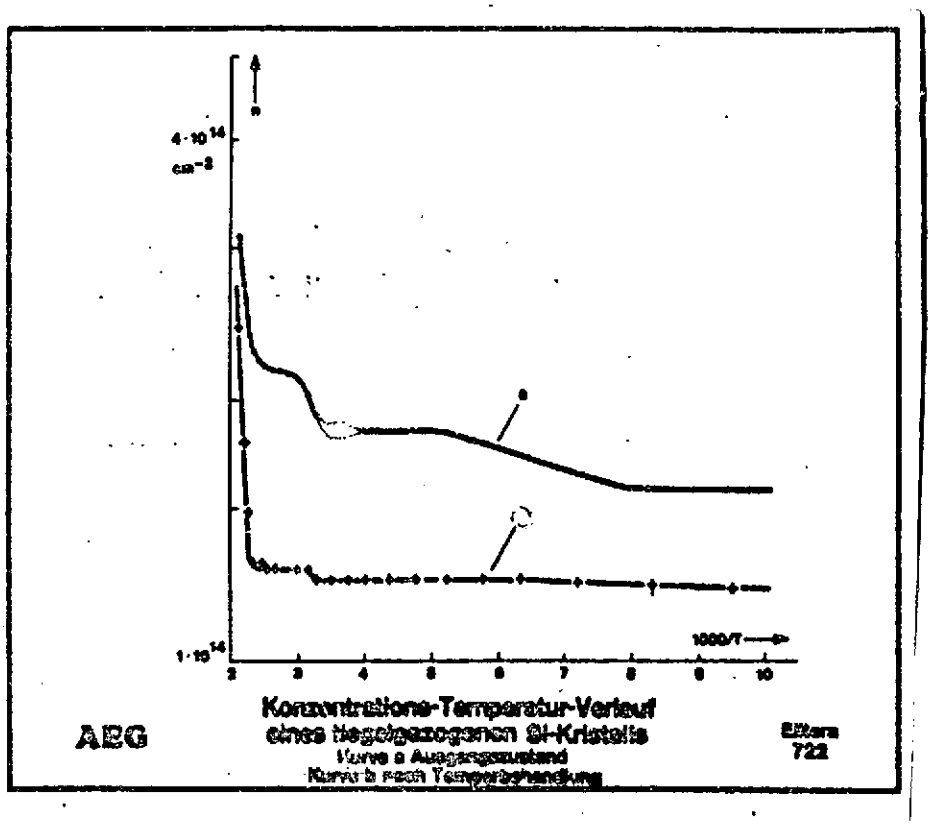


Fig. 27. Concentration-temperature curve of a crucible-pulled Si crystal.

Curve a: initial state
Curve b: after heat treatment

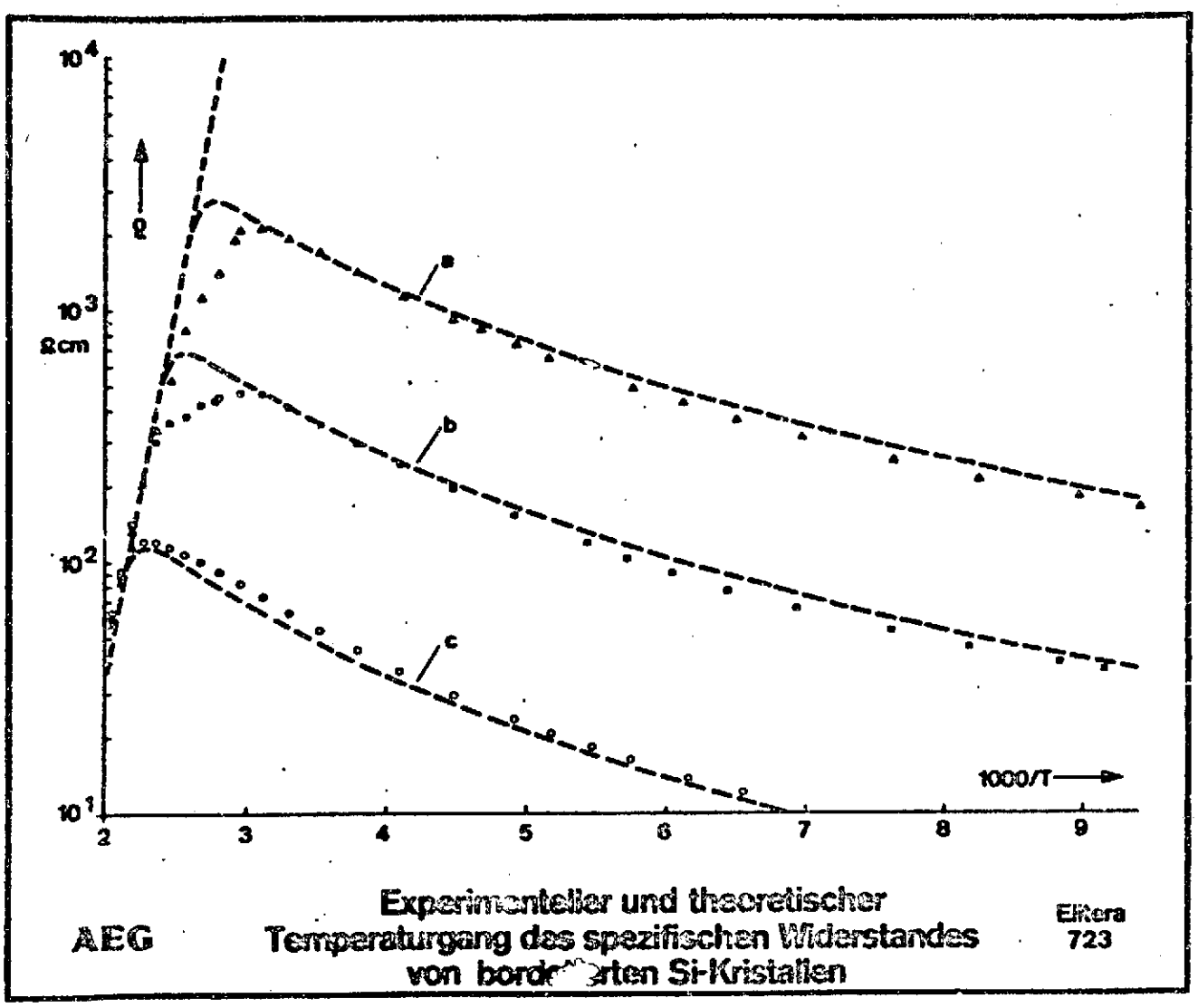


Fig. 28. Experimental and theoretical variation with temperature of resistivity of boron-doped Si crystals.

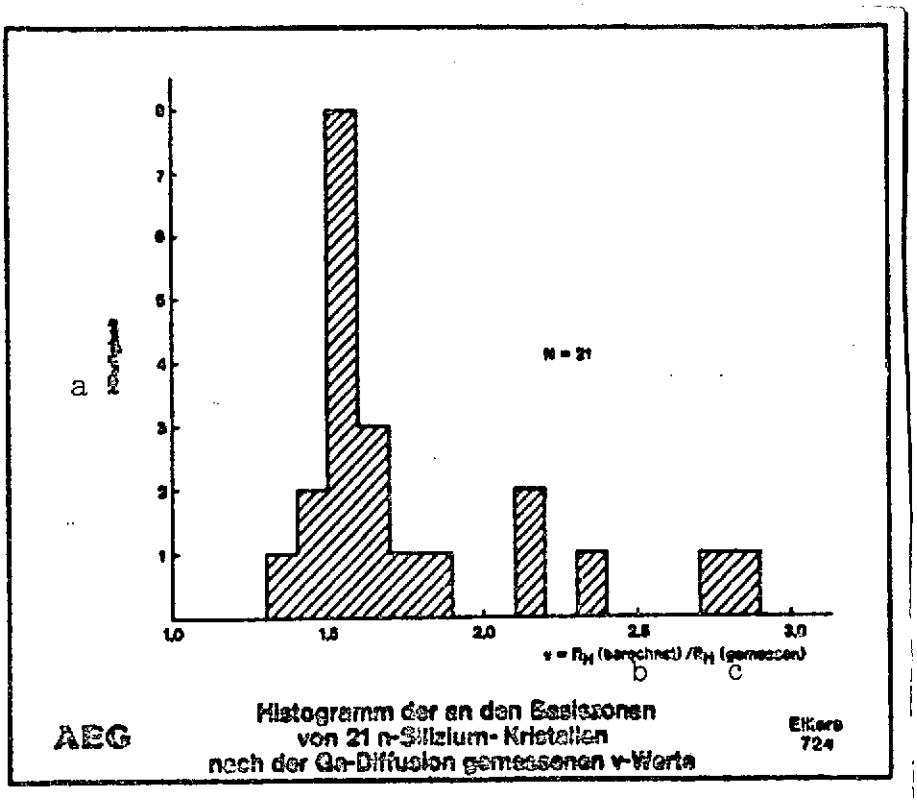


Fig. 29. Histogram of v-values measured for base regions of 21 n-type silicon crystals after Ga diffusion.

Key: a. Frequency
b. Calculated
c. Measured

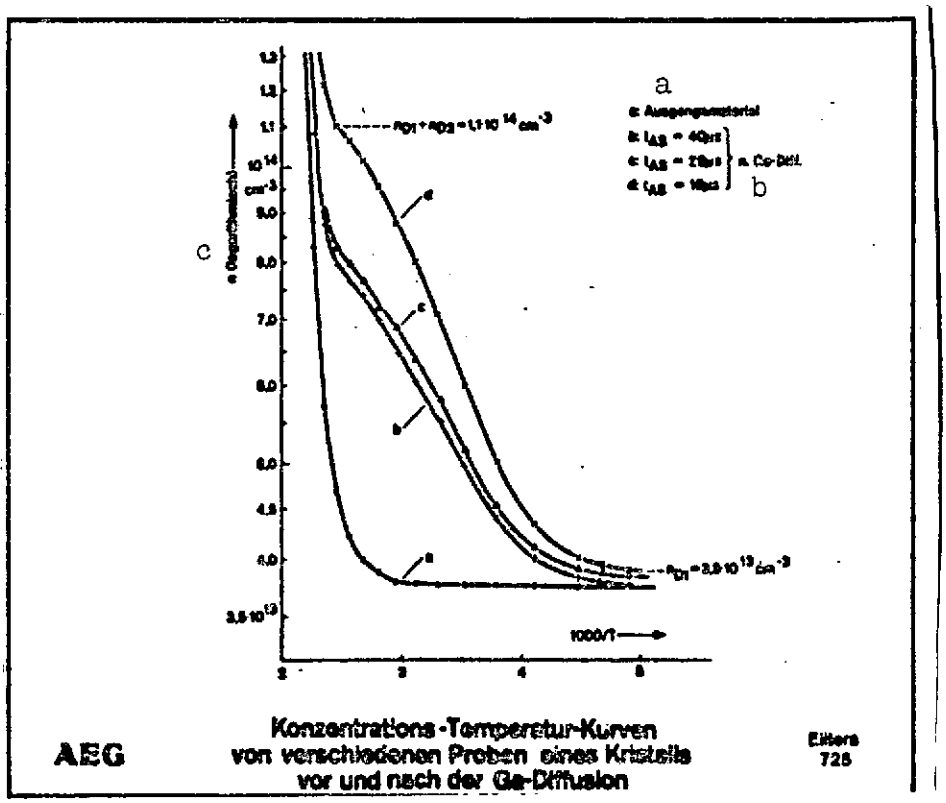


Fig. 30. Concentration-temperature curves of different specimens of a crystal before and after Ga diffusion.

Key: a. Starting material
b. After Ga diffusion
c. Logarithmic

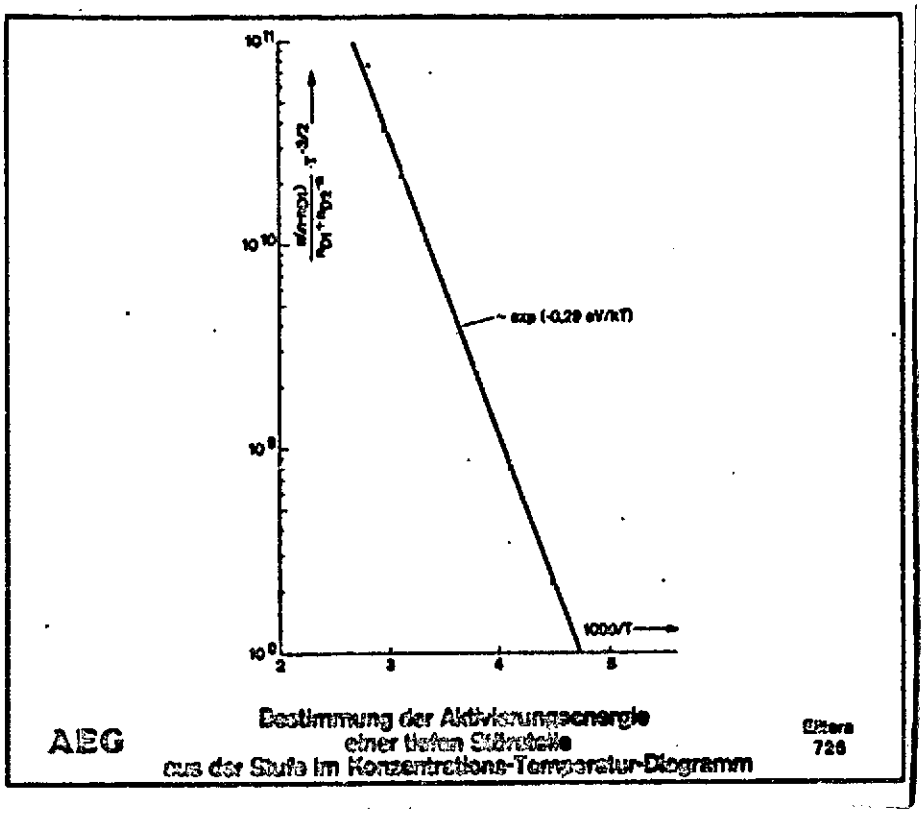


Fig. 31. Determination of activation energy of a deep impurity from the step in the concentration-temperature diagram.

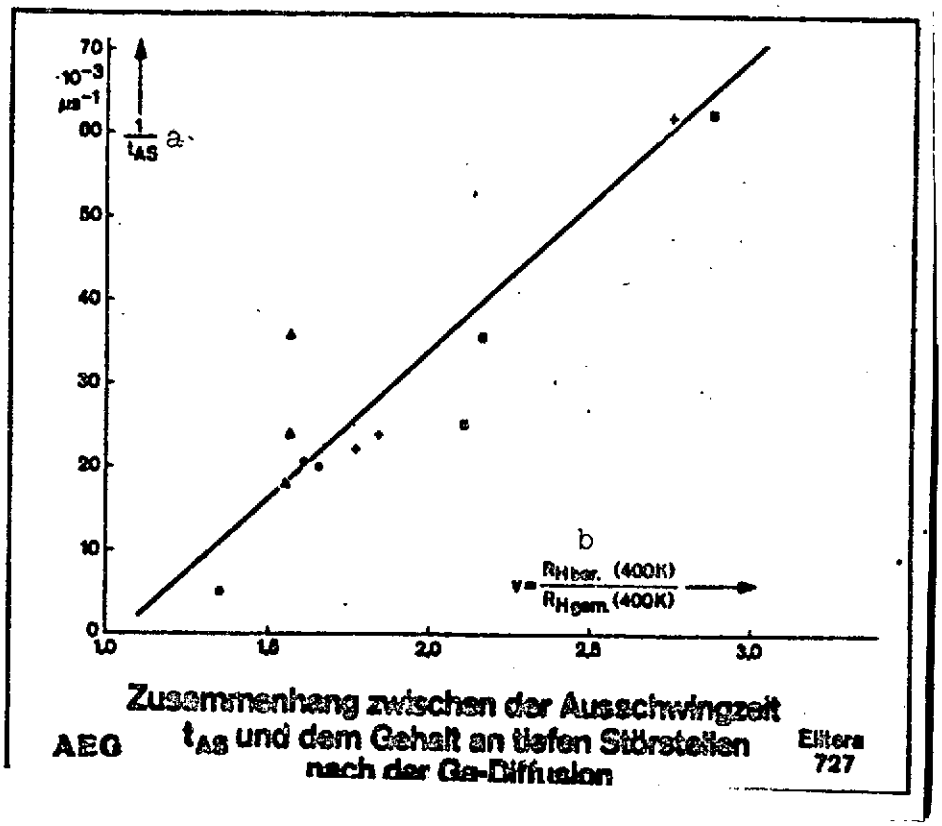


Fig. 32. Relationship between decay time t_d and concentration of deep impurities after Ga diffusion.

Key: a. $1/t_d$
b. $v = R_{Hcalc.}(400\ K)/R_{Hmeas.}(400\ K)$
µs = µsec

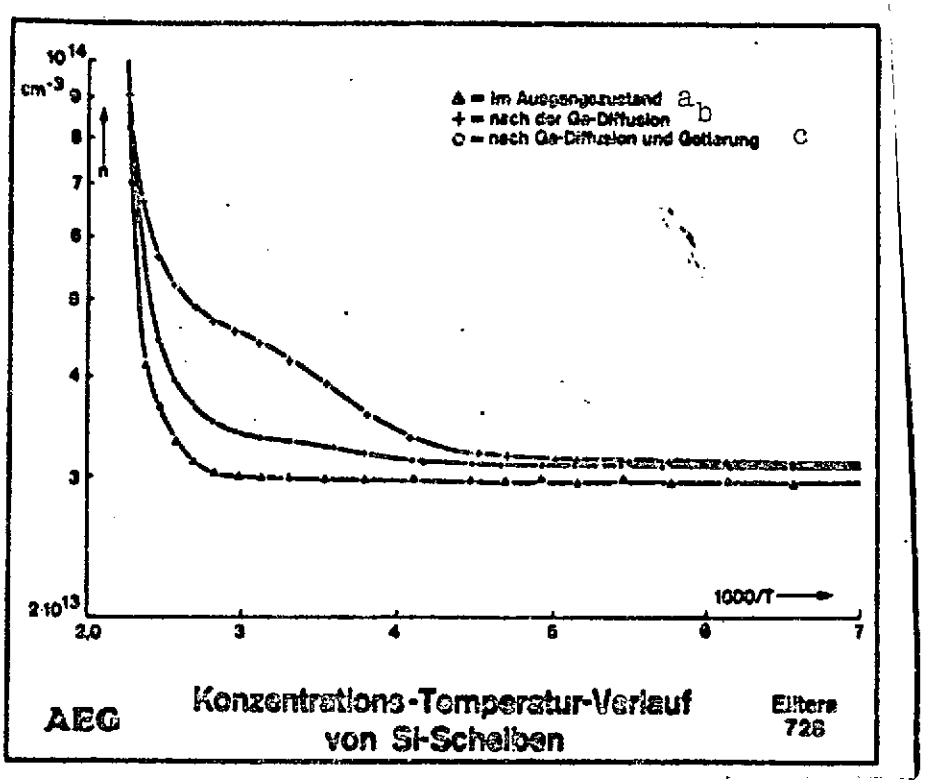


Fig. 33. Concentration-temperature curve of Si slices.

Key: a. In initial state
b. After Ga diffusion
c. After Ga diffusion and gettering.

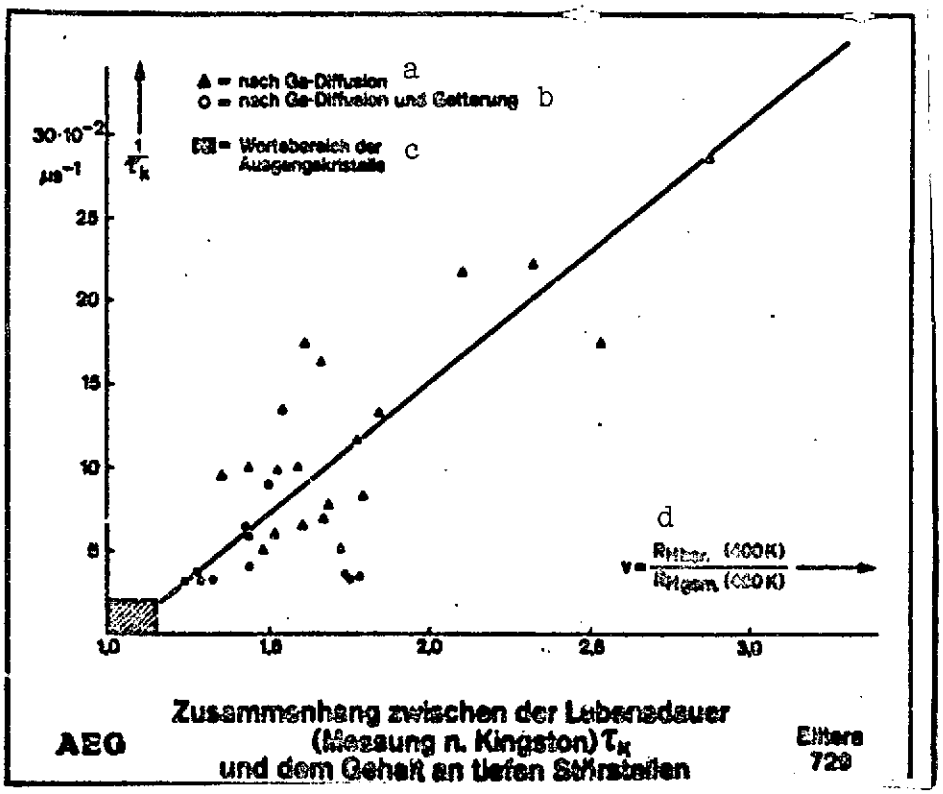


Fig. 34. Relationship between lifetime (Kingston measurement) τ_K and concentration of deep impurities.

Key: a. After Ga diffusion
b. After Ga diffusion and gettering
c. Range of values of starting crystals
d. $R_{Hcalc.}/R_{Hmeas.}$
μs = μsec

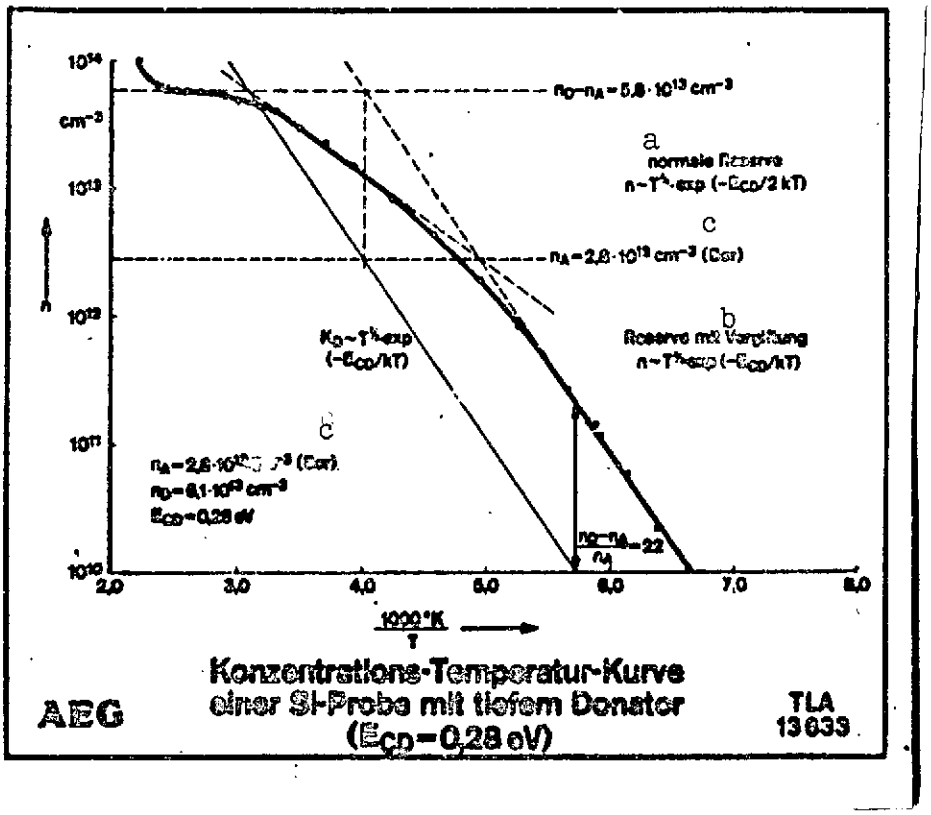


Fig. 35. Concentration-temperature curve of Si specimen with deep donor (E_{CD} = 0.28 eV).

Key: a. Normal reserve
b. Reserve with poisoning
c. Boron

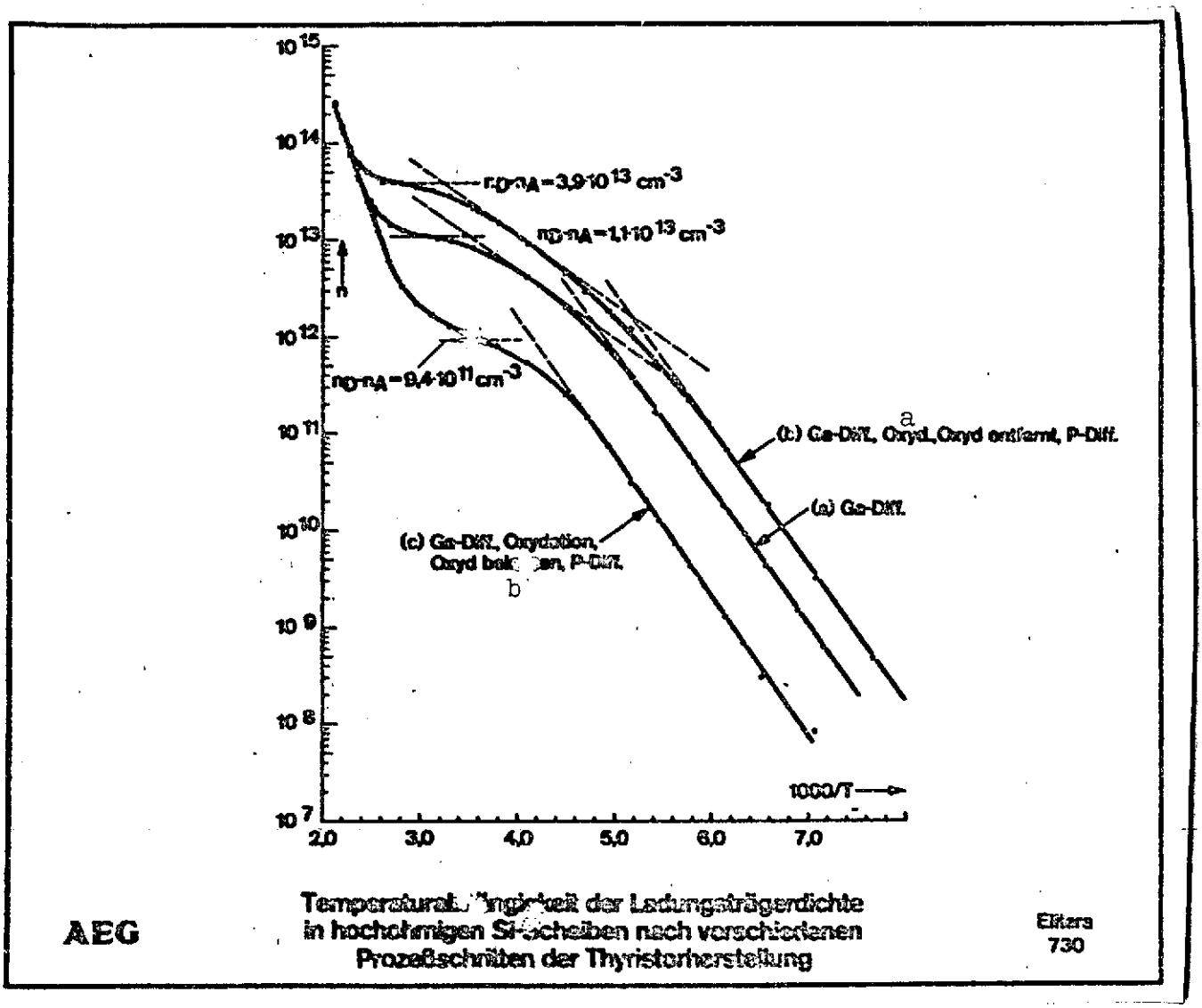


Fig. 36. Temperature dependence of charge carrier density in highly resistive Si slices after various steps in the process of thyristor production.

Key: a. Ga diffusion, oxidation, oxide removed, P diffusion
b. Ga diffusion, oxidation, oxide left on, P diffusion

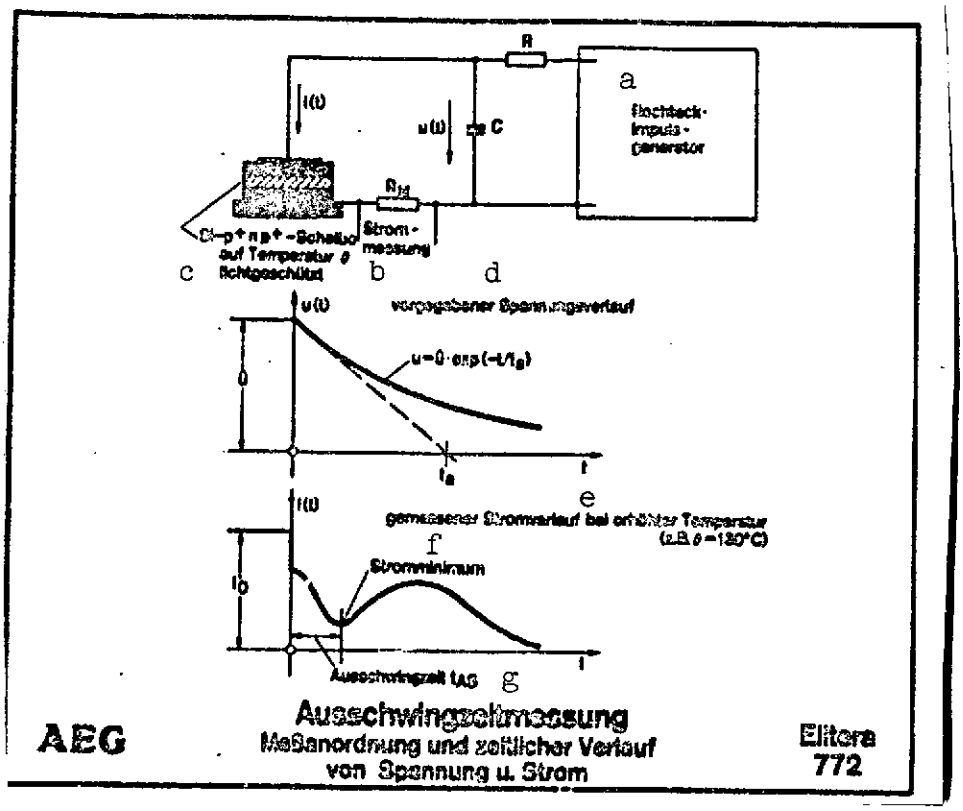


Fig. 37. Decay time measurement measuring apparatus; voltage and current as functions of time.

Key: a. Rectangular pulse generator
b. Current measurement
c. Si p^+np^+ slice at temperature θ, protected from light
d. Applied voltage curve
e. Measured current curve at elevated temperature (e.g. θ = 130°C)
f. Current minimum
g. Decay time t_d

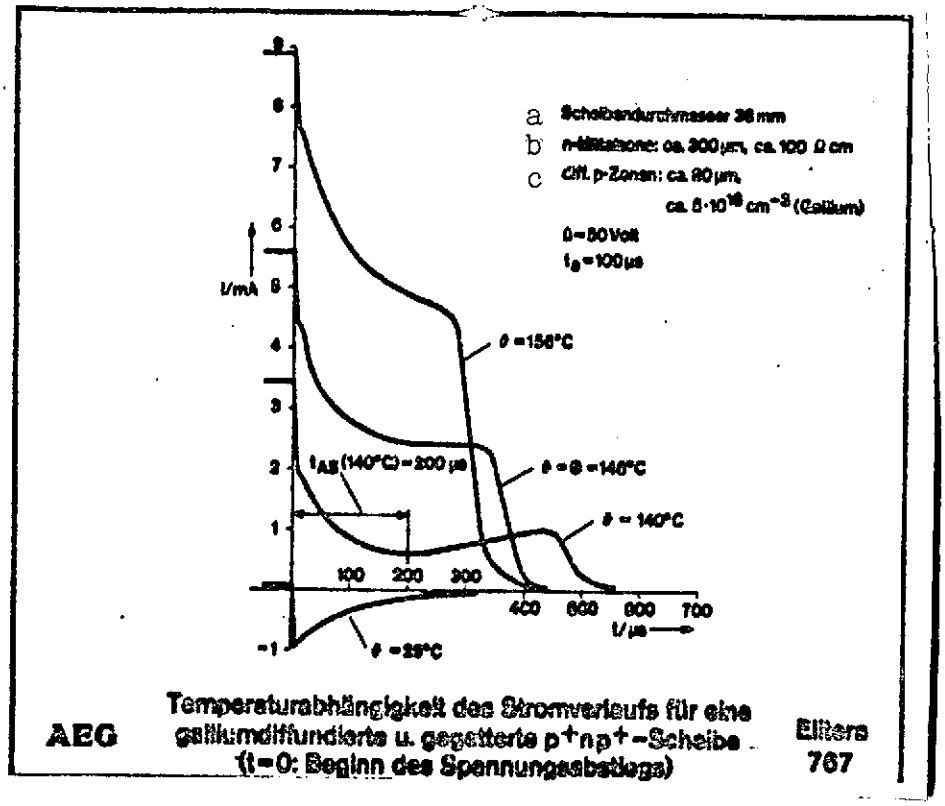


Fig. 38. Temperature dependence of current curve for a gallium-diffused and gettered p^+np^+ slice (t = 0: beginning of voltage drop).

Key: a. Slice diameter
b. Middle n-type region
c. Diff. p-type regions
µs = µsec

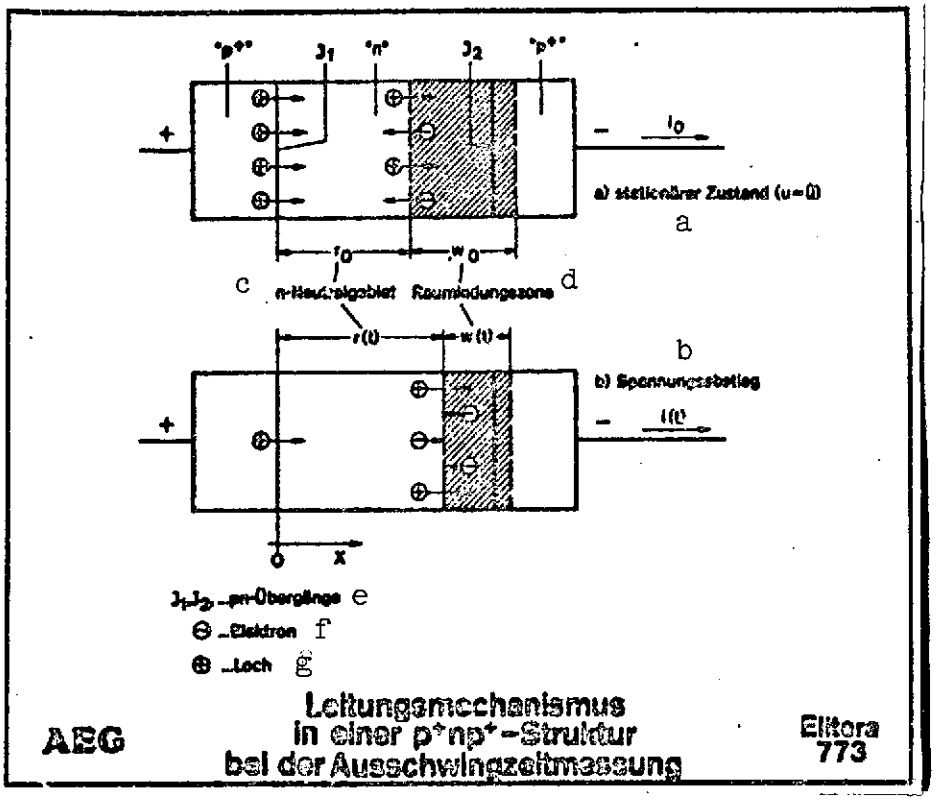


Fig. 39. Conduction mechanism in a p^+np^+ structure in decay time measurement.

Key: a. Steady state
b. Voltage drop
c. n-type neutral region
d. Space charge region
e. p-n junctions
f. Electron
g. Hole

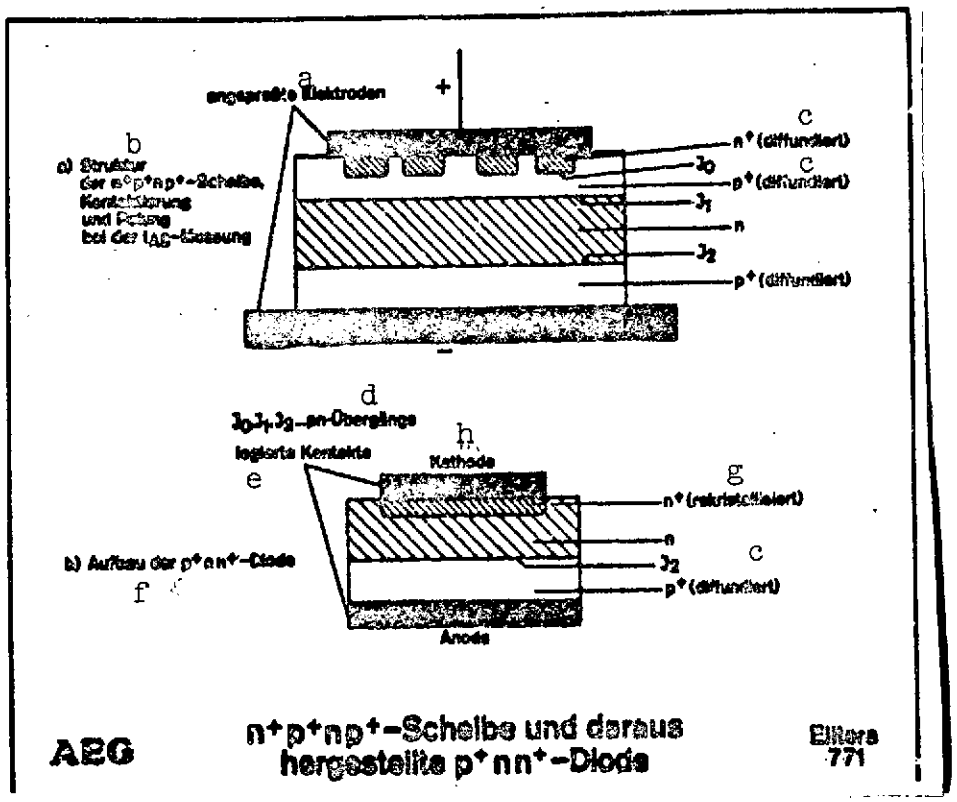


Fig. 40. $n^+p^+np^+$ slice and p^+nn^+ diode prepared from it.

Key:
a. Pressed-on electrodes
b. Structure of $n^+p^+np^+$ slice, contacting and polarity in t_d measurement
c. Diffused
d. p-n junctions
e. Alloyed contacts
f. Structure of p^+nn^+ diode
g. Recrystallized
h. Cathode

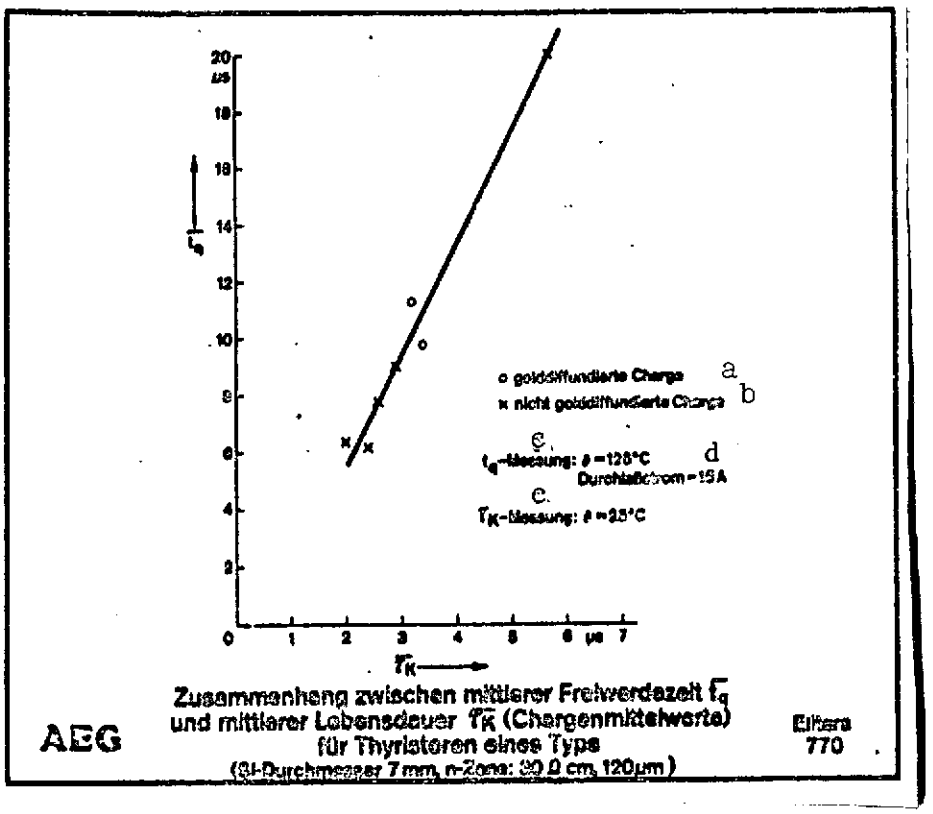


Fig. 41. Relationship between mean recovery time $\overline{t_q}$ and mean lifetime τ_K (mean value for batch) for thyristors of a specific type (Si diameter 7 mm, n-type region: 30 Ωcm, 120 µm).

Key: a. Gold-diffused batch
b. Non-gold-diffused batch
c. Measurement
d. Forward current
µs = µsec

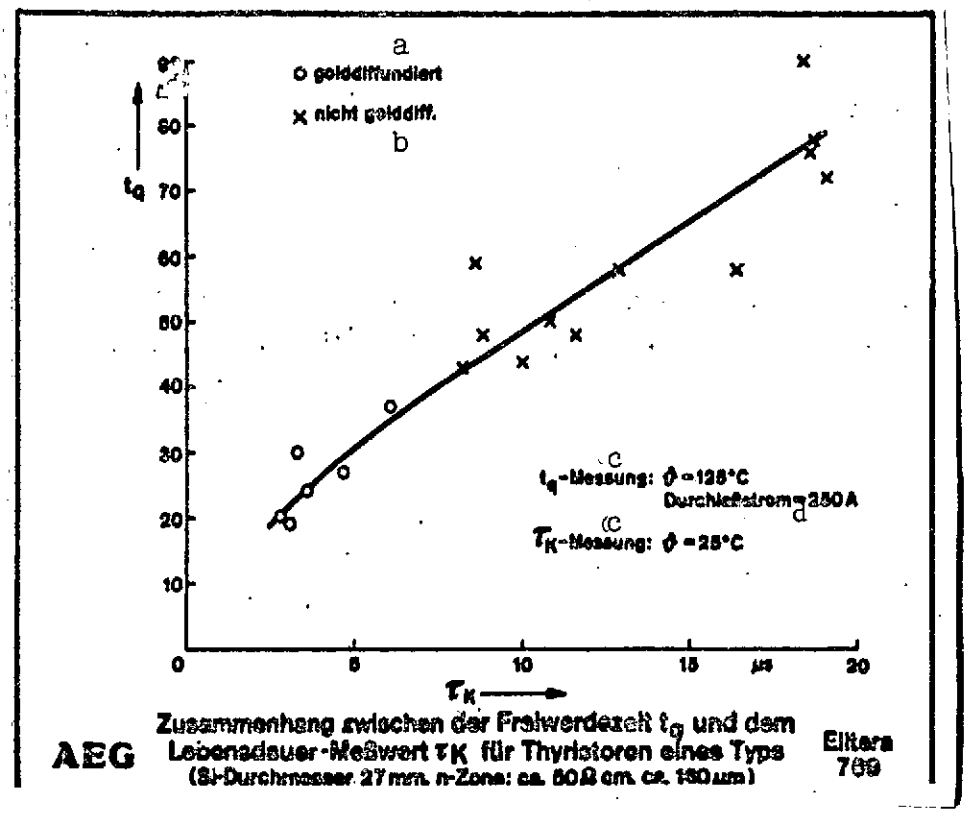


Fig. 42. Relationship between recovery time t_q and mean lifetime τ_K for thyristors of a single type (Si diameter 27 mm, n-type region: about 50 Ωcm, about 160 µm).

Key: a. Gold-diffused
b. Non-gold-diffused
c. Measurement
d. Forward current
µs = µsec

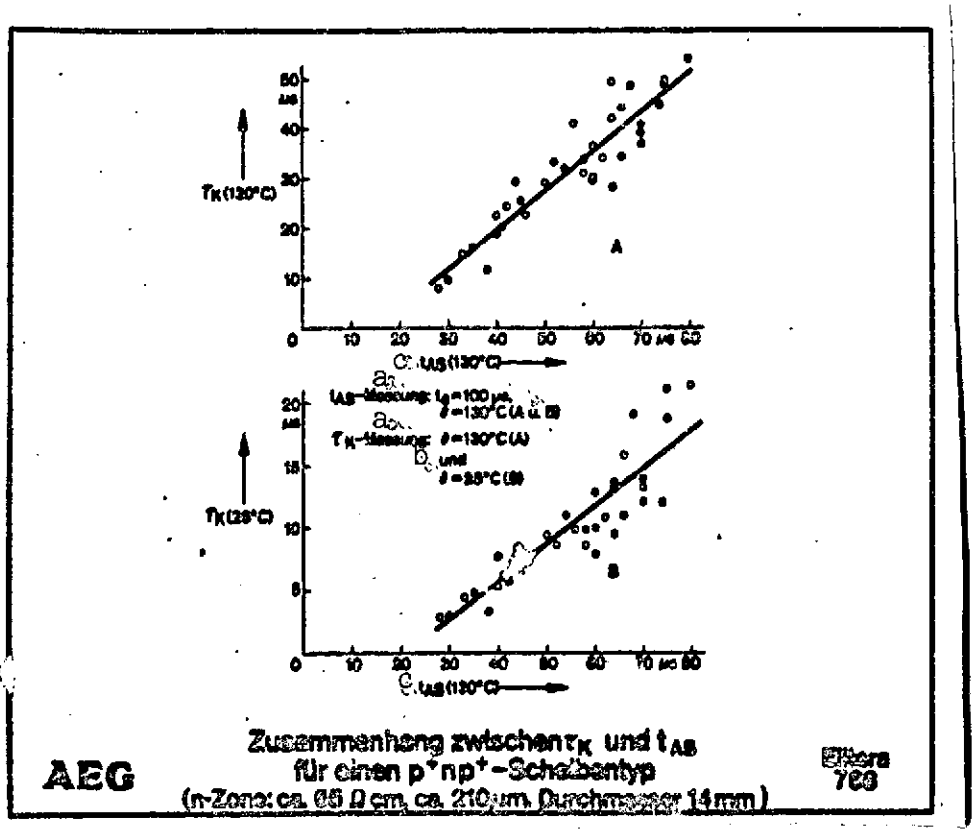


Fig. 43. Relationship between τ_K and t_d for a p^+np^+ slice type (n-type region: about 65 Ωcm, about 210 µm, diameter 14 mm).

Key: a. Measurement
b. And
c. t_d
µs = µsec

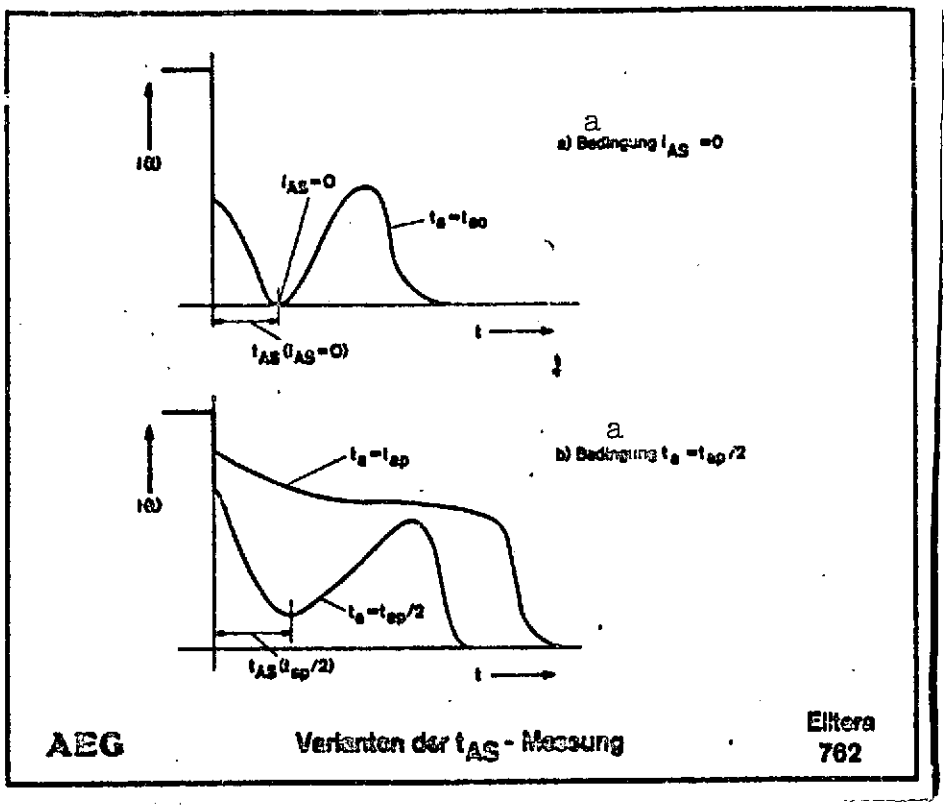


Fig. 44. Variants of t_d measurement.

Key: a. Condition
$t_{AS} = t_d$

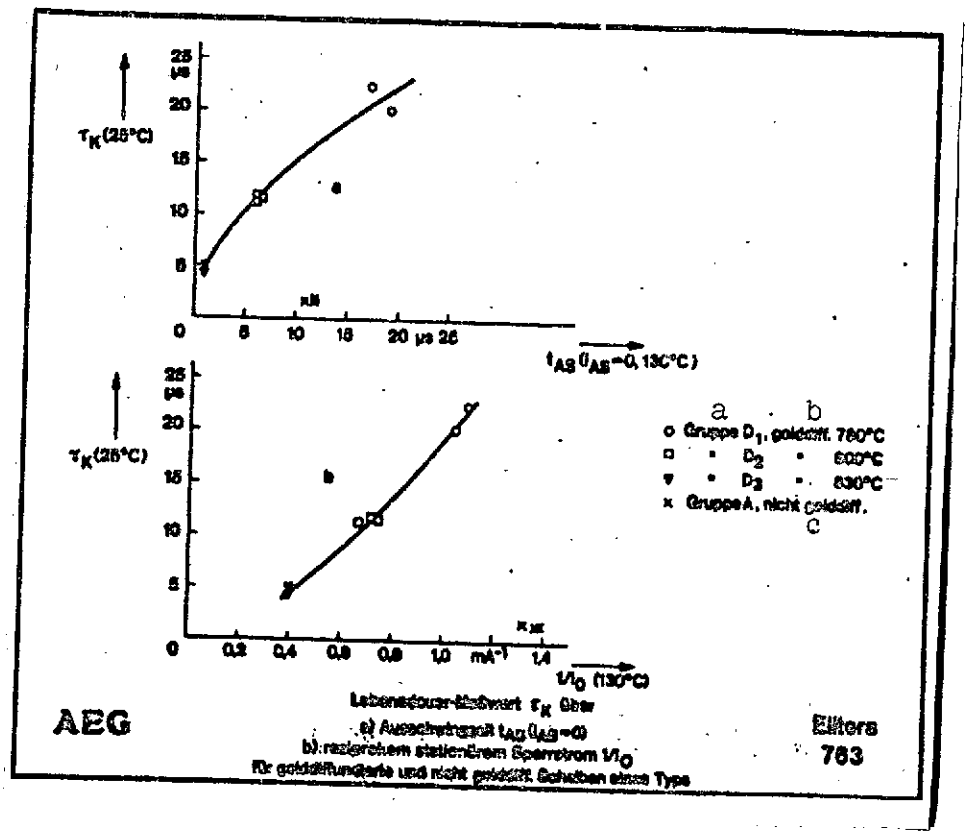


Fig. 45. Measured lifetime τ_K as a function of
a) decay time t_d ($i_d = 0$)
b) reciprocal $1/i_o$ of steady-state reverse current
for gold-diffused and non-gold diffused slices of a single type.

Key: a. Group
b. Gold-diffused
c. Non-gold-diffused
$t_{AS} = t_d$

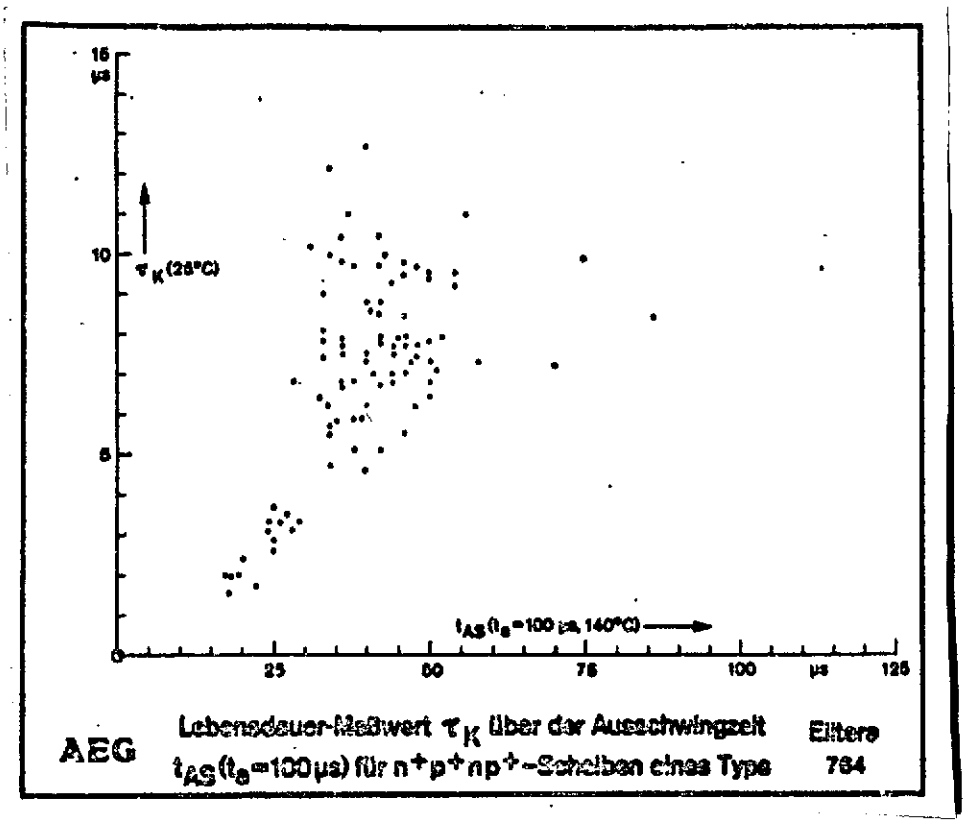


Fig. 46a. Measured lifetime τ_K as a function of decay time t_d(t_a = 100 µsec) for $n^+p^+np^+$ slices of a single type.

Key: µs = µsec
$t_{AS} = t_d$

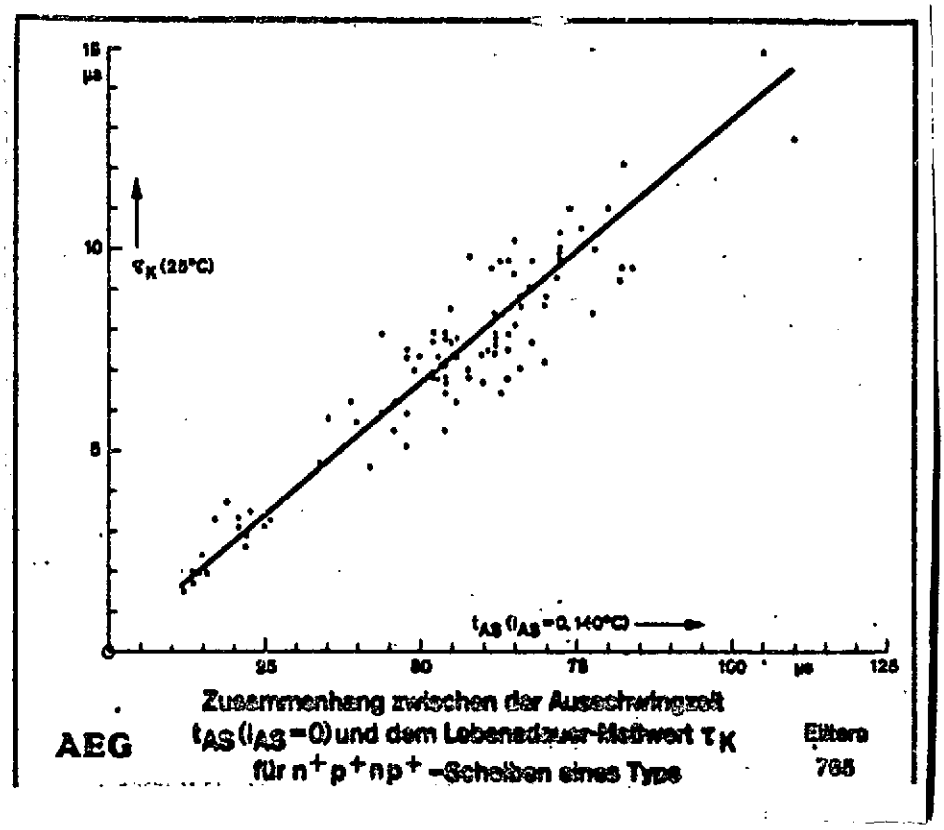


Fig. 46b. Relationship between decay time $t_d(i_d = 0)$ and measured lifetime τ_K for $n^+p^+np^+$ slices of a single type.

Key: µs = µsec
$t_{AS} = t_d$

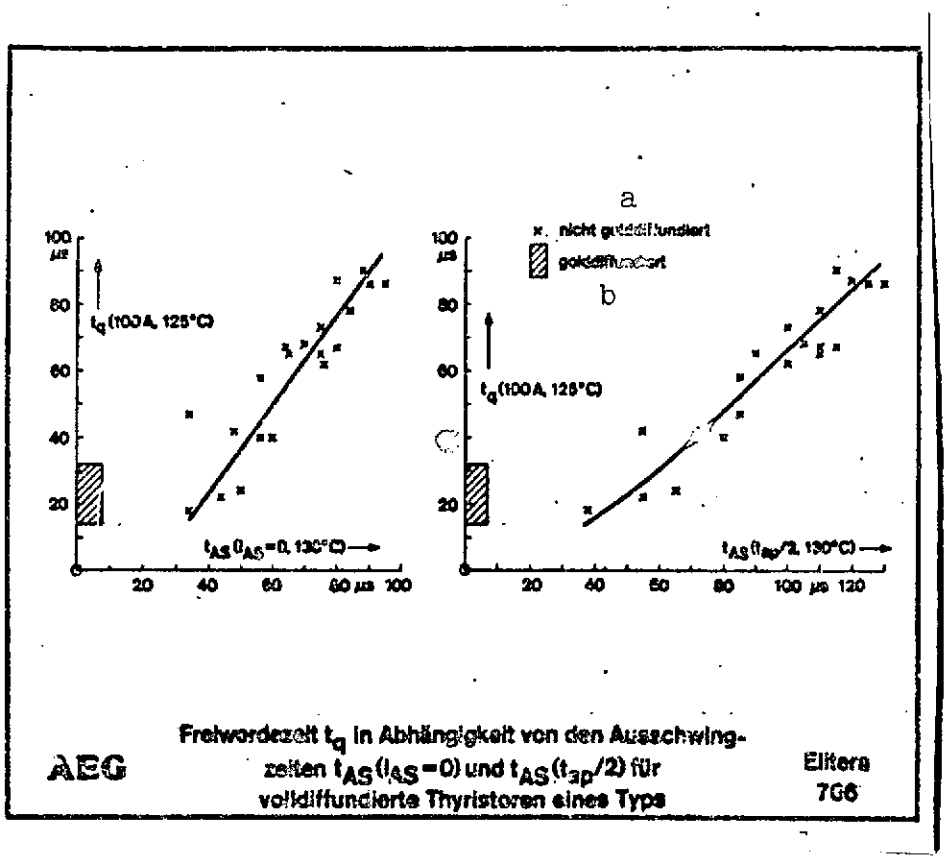


Fig. 47. Recovery time t_q as a function of decay times $t_d(i_d = 0)$ and $t_d(t_{ap}/2)$ for fully diffused thyristors of a single type.

Key: a. Non-gold-diffused
b. Gold-diffused
µs = µsec
$t_{AS} = t_d$

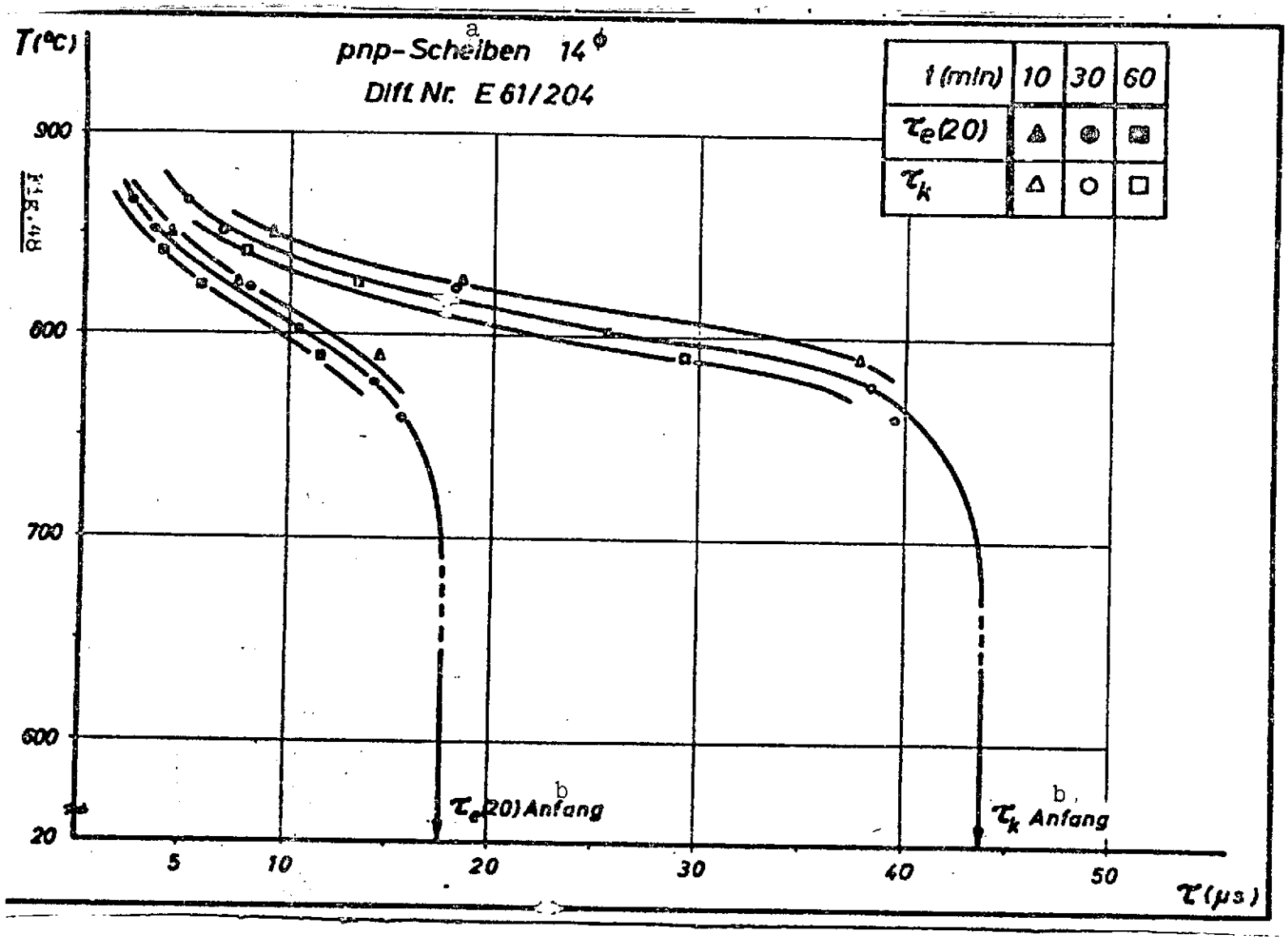


Fig. 48. Lifetime τ in relation to diffusion time and temperature.

Key: a. p-n-p slices; b. start; ϕ = diameter.

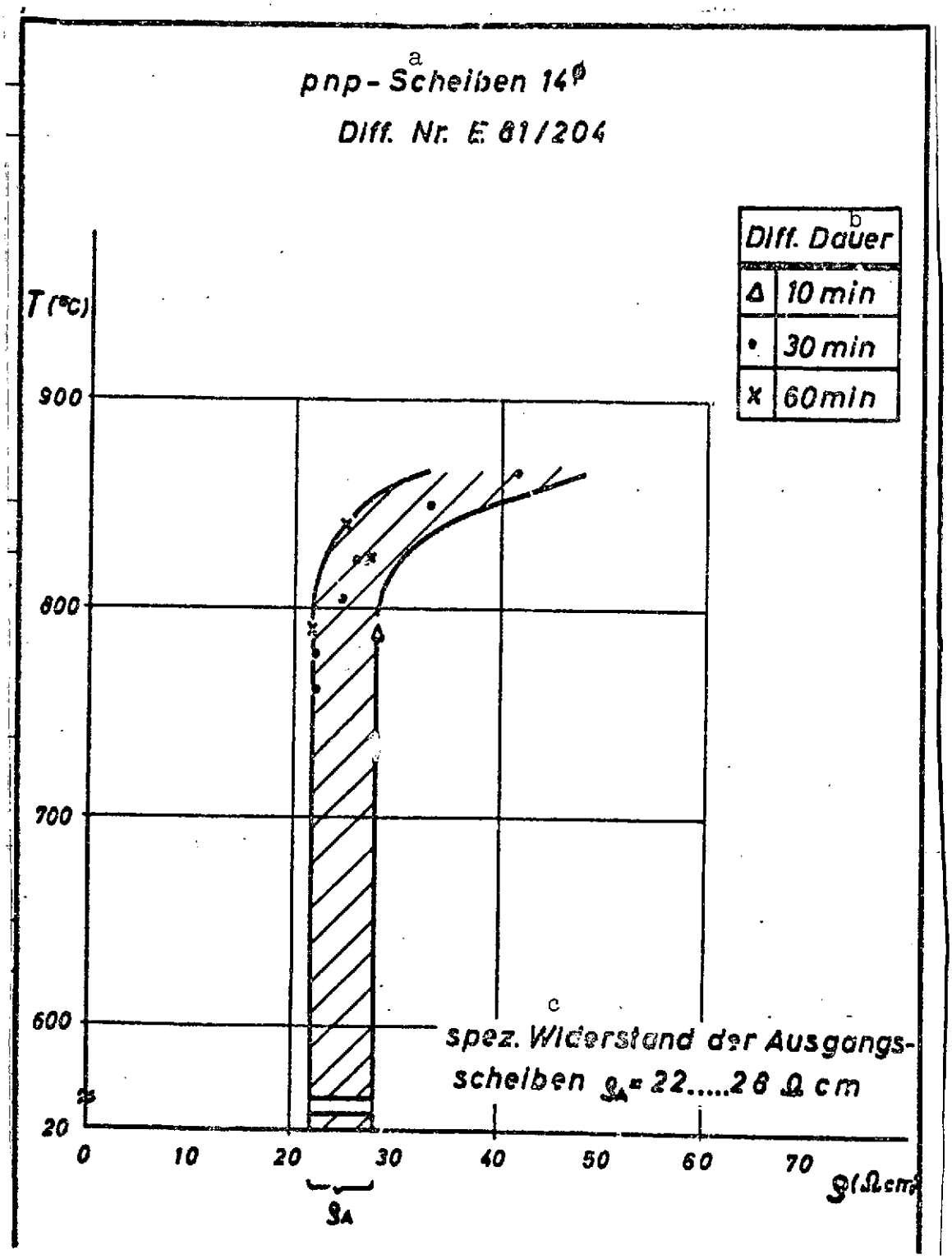


Fig. 49. Resistivity ρ as function of diffusion temperature T.

Key: a. p-n-p slices
b. Duration of diffusion
c. Resistivity of starting slices
ø = diameter

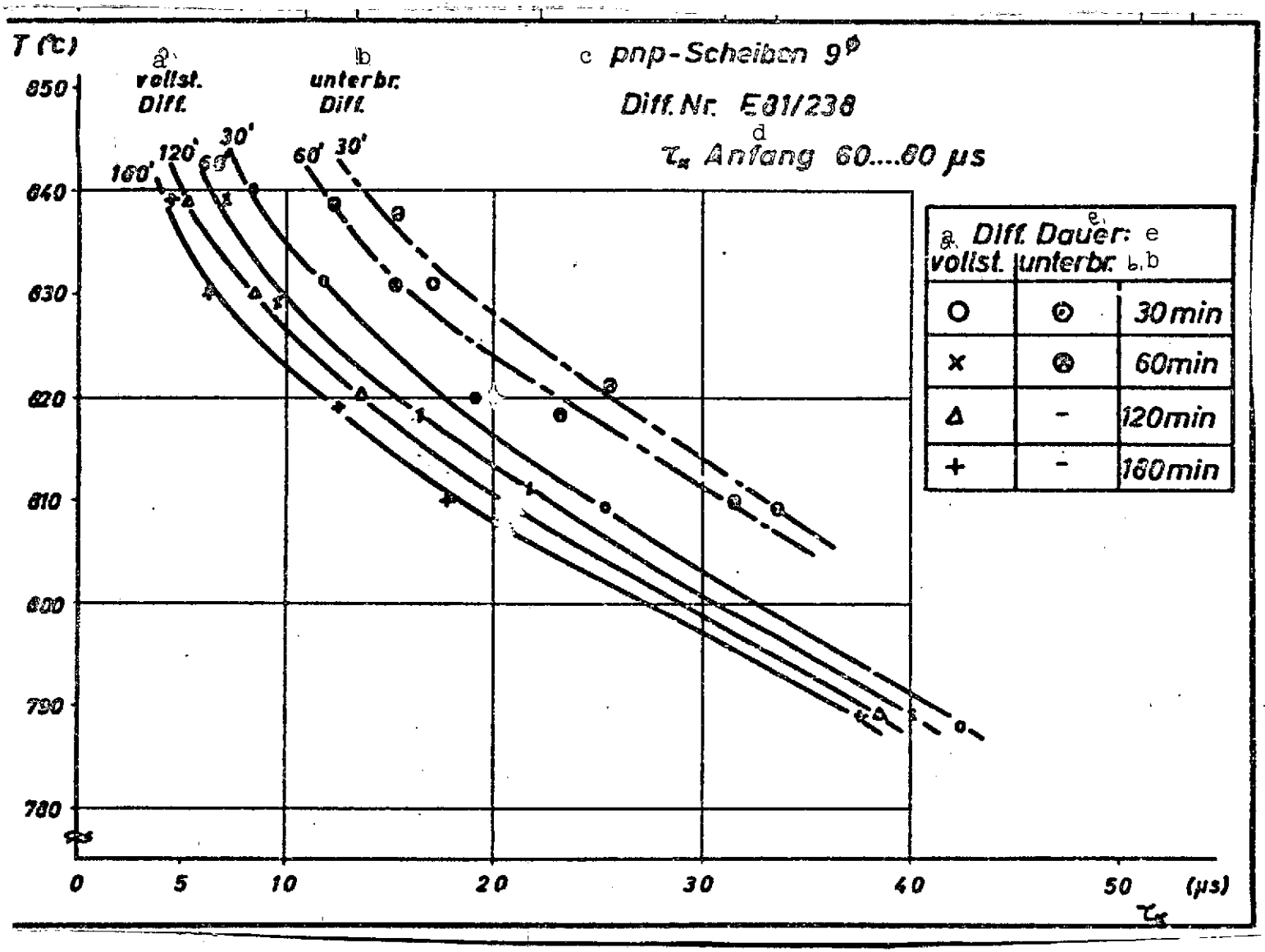


Fig. 50. Lifetime τ in relation to diffusion temperature, diffusion time, and preliminary treatment.

Key: a. Complete diffusion; b. interrupted diffusion; c. p-n-p slices; d. start; e. duration of diffusion; µs = µsec; ø = diameter.

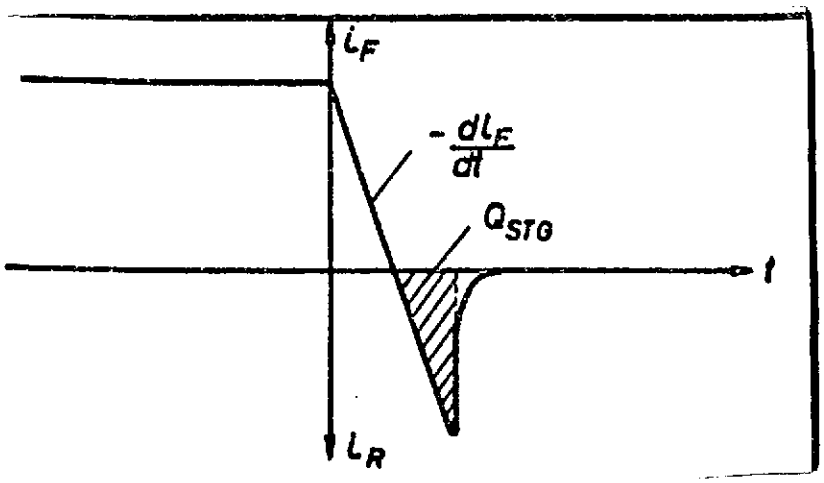


Fig. 51. Definition of storage charge Q_{STG}.

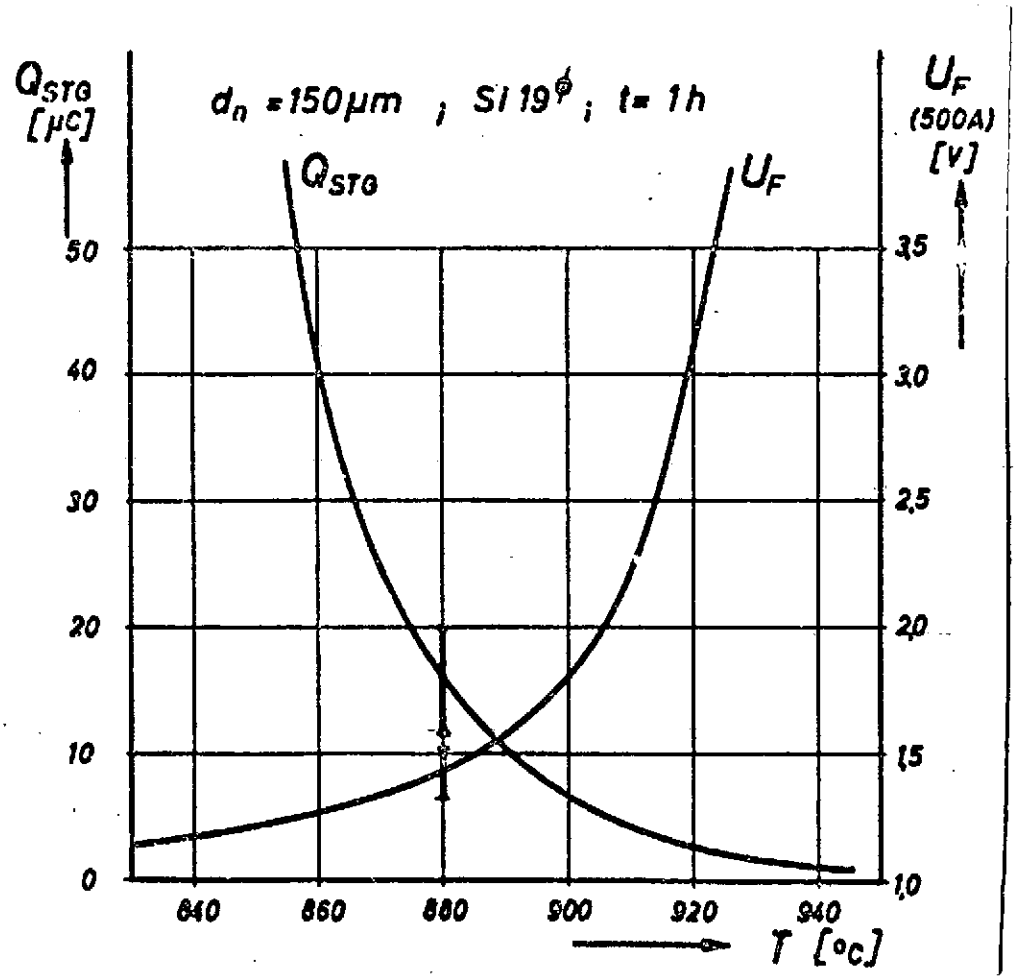


Fig. 52. Q_{STG} and U_F as functions of T.
Key: ϕ = diameter; h = hour.

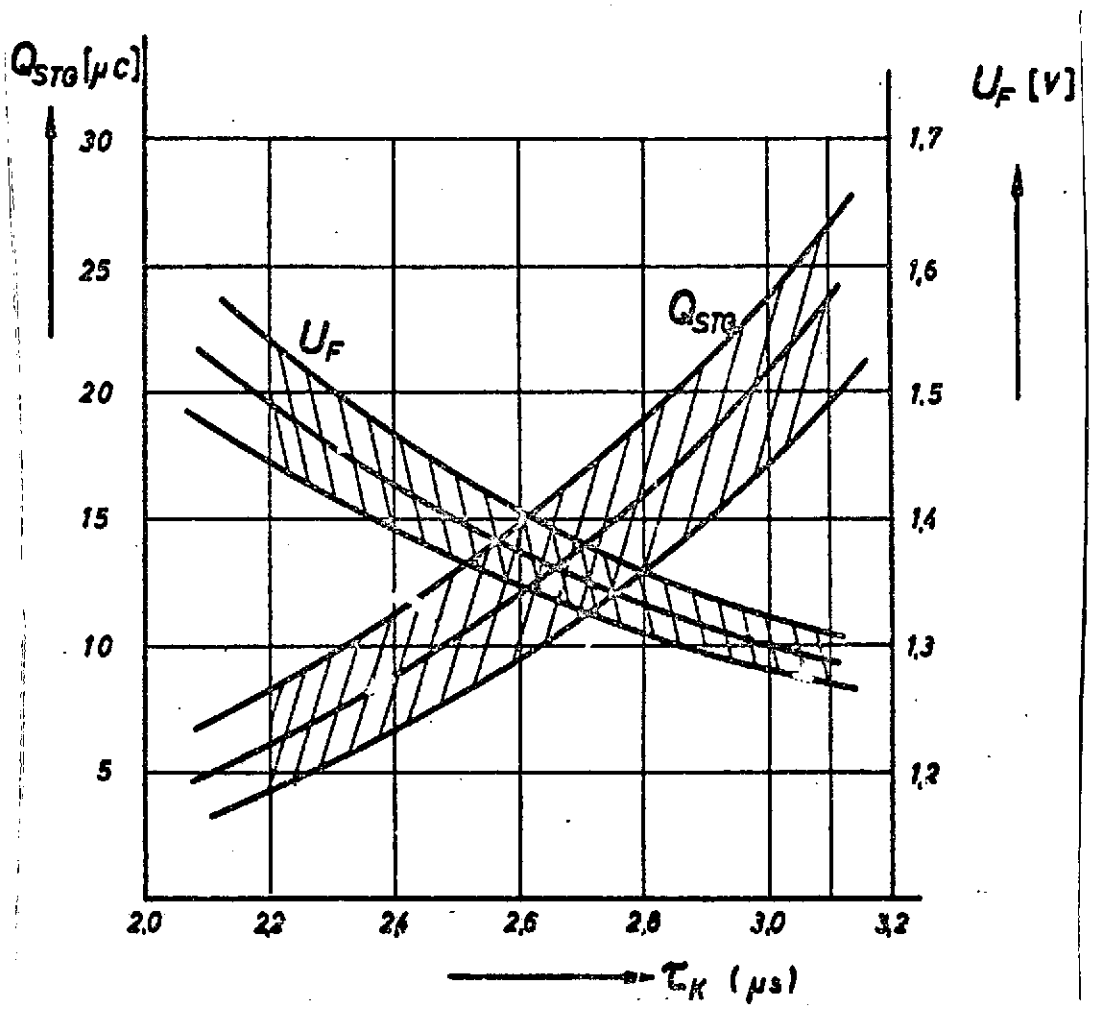


Fig. 53. Q_{STG} and U_F as functions of lifetime τ_K.

Key: μs = μsec

TABLE 1.

/183

Diode 1, $W_n = 900\,\mu m$
Gold -diffused

	I_F (mA)	20	10	5	2	1
$I_R = 0$	τ_p^{PI} (µs)	0.65		0.6		0.2
$I_R = 0{,}1$ mA	t_s (µs)	9.2	8.1	7.1	6.0	5.5
	τ_p^S (µs)	2.3	2.4	2.6	3.0	3.8
$I_R = 1{,}0$ mA	t_s (µs)	3.4	2.4	1.7	1.0	0.6
	τ_p^S (µs)	1.7	1.7	1.8	2.1	2.6

Diode 2, $W_n = 800\,\mu m$

20	10	5	2	1
3.0		2.5		2.0
24	20	17.0	13	10.4
6.0	6.0	6.2	6.6	7.2
10.4	7.9	5.4	2.9	1.8
5.3	5.5	5.7	5.8	7.7

Diode 3, $W_n = 3$ mm
Gold -diffused

	I_F (mA)	20	10	5	2	1
$I_R = 0$	τ_p^{PI}	3.0	3	2.5	2.0	1.2
$I_R = 0{,}1$ mA	t_s	23	19	16	11.5	9
	τ_p^S	5.7	5.7	5.9	5.4	6.2
$I_R = 1{,}0$ mA	t_s	12	8	5.4	3.9	1.4
	τ_p^S	6.1	5.9	5.7	5.4	6.1

Diode 4, $W_n = 4$ mm
Gold - diffused

20	10	5	2	1
10		8.5		7.0
60	51	43	34	25
15	15.2	15.8	17.3	17.3
3[illegible]	23.5	15	[illegible]	4.8
15.3	16.4	17.0	17.7	20.8

Diode 5, $W_n = 3$ mm

	I_F	20	10	5	2	1
$I_R = 0$	τ_p^{PI}	20		18		14
$I_R = 0{,}1$ mA	t_s	126	106	84	58	42
	τ_p^S	31.5	32	31.0	29.6	29.2
$I_R = 1{,}0$ mA	t_s	69	50	34	16	8
	τ_p^S	35	34.6	36.0	32.2	34.8

Diode 6, $W_n = 3$ mm

20	10	5	2	1
85		70		50
455	370	300	205	142
116	110	110	105	99
260	198	124	56	26
132	130	132	117	116

TABLE 2.

/184

Diode	τ_p^{PI} (μs) (20mA)	τ_p^{S} (μs) I_R = 0.1 mA	τ_p^{S} (μs) I_R = 1.0 mA
1	0.9	2.8	2.0
2	3.0	6.4	6.0
3	3.0	5.9	5.8
4	10.0	16.1	17.6
5	20.0	30.7	34.5
6	85.0	108	125.4

TABLE 3. VALUES OF τ BEFORE AND AFTER GOLD DIFFUSION. /185

Starting silicon - Producer and crystal number	Gold diffusion conditions			
	785°C/1h		805°C/1h	
	τ_p'' (μs)	$\tau_{p(t)}$ (μs)	τ_p'' (μs)	$\tau_{(t)}$ (μs)
Wacker/AEG	90	4,9	140	2,9
4K/ZN 325	16,5	6,8		
Haldor Topsøl (DK)	115	12,2	112	4,8
37-06-01	160	7,5	21	4,5
Wacker Chemitronic 300400/9	280	14,4	145	7,5
Komatsu / Japan FV 36497	16	6,9		
Texas Inst./ USA Lopex 35442	15,2	13,8	16,0	10,2

Dislocation density in all crystals $\approx 10^4$ cm^{-2}, except for Lopex 35442 ≈ 0.

TABLE 4.

t (h)	0	1	2	3	5	7	12	28	93	180
Diode 1 Tp(t) (µs)	90.5	4.9	3.4	3.0	2.6	2.2	2.1	1.8	1.5	1.5
Diode 2 Tp(t) (µs)	115	11.6	7.5	6.6	5.4	4.1	3.4	2.5	1.5	1.5
Diode 3 Tp(t) (µs)	160	7.6	5.9	5.0	4.1	3.4	2.8	2.3	1.6	1.5

Diode 6 (Lopex 35442) $\tau_p'' = 15{,}2\ \mu s$, $\tau_{min} = 1{,}5\ \mu s$

/187

Diffusion time t(h)	$\tau_p(t)$ (μs)	$\frac{1}{\tau_p(t)}$ $(\mu s)^{-1}$	$\frac{1}{\tau_p(t)} - \frac{1}{\tau''}$ $(\mu s)^{-1}$	$\frac{\tau_{min}}{\tau_p}$
3	10.8	0.0925	0.0265	0.04
5	9.8	0.102	0,036	0.054
7	8.3	0.120	0.054	0.081
12	7.0	0.143	0.077	0.115
28	5.4	0,185	0.119	0.18
92	3.7	0.27	0.20	0.30
180	2.8	0.357	0.291	0.44

Diode 7 (Komatsu 34973), $\tau_p'' = 16{,}0\ \mu s$, $\tau_{min} = 1{,}5\ \mu s$

Diffusion time t(h)	$\tau_p(t)$ (μs)	$\frac{1}{\tau_p(t)}$ $(\mu s)^{-1}$	$\frac{1}{\tau_p(t)} - \frac{1}{\tau''}$ $(\mu s)^{-1}$	$\frac{\tau_{min}}{\tau_p}$
3	6.8	0.147	0.085	0.13
5	5.9	0.17	0,11	0.165
7	5.0	0.20	0.14	0.21
12	4.2	0.24	0.18	0.27
28	3.2	0.31	0.25	0.375
92	2.0	0.5	0.44	0.66
180	1.5	0.67	0.61	0.91

TABLE 6. /188

Diode Nr.	Crystal designation	τ_p° (µs)	τ_p (t) after 1h/820 °C (µs)	τ_p (t) after further 90h/780 °C (µs)
1	4K/ZN 325	110	2,0	1.5
2	4K/ZN 325	135	1,8	1.4
3	7K/WZN825	14.8	2.7	1.7
4	Lopex 17964	6.0	3.1	2.8
5	6K/WZN733	10.1	3.7	1.6
6	HT. 307061	120	3.0	1.5

TABLE 7.

/189

Gettering Temperature	Concentrations per cm^3 in phosphorus silicate layers		Residual concentrations per cm^3	
	Au	Cu	Au	Cu
800°C	2.2×10^{16}	4×10^{18}	2×10^{12}	$< 10^{12}$
900°C	1.5×10^{16}	1.0×10^{18}	4.1×10^{11}	$< 10^{12}$
1000°C	$1.\times10^{16}$	3.0×10^{18}	1.4×10^{11}	$< 10^{12}$
1100°C	4.0×10^{15}	1.5×10^{18}	5.0×10^{10}	$< 10^{12}$
1200°C	1.0×10^{16}	1.2×10^{18}	$< 2\times10^{10}$	$< 10^{12}$

TABLE 8. GOLD DIFFUSION CONDITIONS, 20 HOURS AT 900°C GOLD CONCENTRATION = $3.25 \cdot 10^{15}$ cm^{-3} (INITIAL); GETTERING OR DIFFUSION TEMPERATURE = 1160°C. /190

Gettering or Diffusion	Time (hour)	Penetration depth (μm)	Surface concentration per cm^3	Residual Au concentration per cm^3
Phosphorus	1¾	85	10^{21}	3.6×10^{12}
				1.4×10^{11}
Arsenic	20	65	8×10^{20}	1.24×10^{13}
				1.27×10^{13}
Boron	1½	100	4×10^{20}	3.44×10^{15}
				3.37×10^{15}
Gallium	2	——	——	4.2×10^{11}
				8.9×10^{10}

TABLE 10. INFLUENCE OF GETTERING AND GOLD DIFFUSION ON LIFETIME AND DECAY TIME OF GALLIUM-DIFFUSED SLICES.

Group	N	τ_K 25°C /µs/	t_d (t_a=100µs) 130 °C /µs/	t_d (i_d=0) 130 °C /µs/	Θ (t_a=100µs) /C°/	i_o 130 °C /mA/
A	3	1.7-1.8 - 1.9	15 - 16 - 17	11 - 12 - 12	142-143-144	0.73-0.74-0.75
B	6	43.5-46.4-54.5	65 - 98 -200	95 -131 -175	137-144-147	0.62-0.81-1.25
C	2	53.4-57.5-61.6	105-120 -135	235 -235 -235	141-142-142	1.40-1.53-1.65
D_1	2	20.2-21.3-22.4	28- 31 - 34	17 - 18 - 19	137-138-138	0.92-0.94-0.96
D_2	3	11.3-11.5-11.8	not measurable	6 - 6 - 6	120-122-123	1.35-1.42-1.50
D_3	2	4.2- 4.6- 4.9	not measurable	1 - 1 - 1	106-107-107	2.50-2.53-2.55

N = number of p^+np^+ slices studied per group

Further explanations, see Section [6].5.3.

Commercial proprietary rights (invention patents and patented designs) from the development program: /193

1) Title:

Procedure for determining the carrier lifetime in semiconductor component slices with at least three regions of alternating conduction type.

The invention concerns a procedure for determining the carrier lifetime in the interior of semiconductor component slices with at least three regions of alternating conduction type. After an edge etching treatment, the component slices are placed between two contact electrodes. They are protected against light and held at a temperature between 50° and 160°C, while a d.c. voltage is applied whose polarity is such that only one p-n junction of each slice is polarized in the reverse direction. The voltage is held constant for 1 msec and then decreased exponentially. Current is measured as a function of time. The time interval until the current minimum is a measure for the carrier lifetime.

Document No.:
P 20 37 089.2

Application Date:
July 27, 1970

Disclosure Date:
February 10, 1972

Applicant:
Licentia Patent-Verwaltungs-GmbH
6000 Frankfurt

Inventor:
Borchert, Edgar
4785 Belecke
Gertkemper, Paul
4771 Sichtigvor

2) Title:

Procedure for doping a semiconductor

The invention concerns a procedure for doping a semiconductor by indiffusion of gallium. The gallium diffused inward from a liquid, gallium-containing phase on the surface of the semiconductor.

Document:
P 20 29 221.1

Application Date:
June 13, 1970

Disclosure Date:
December 16, 1971

Applicant:
same as before

Inventors:
J.L. Lambert
Dr. G. Köhl

31 Title:

Procedure for producing a semiconductor component

The invention concerns a procedure for producing a semiconductor component, the semiconductor of which contains a predetermined concentration of deep impurities acting as recombination centers. After the production of the semiconductor and if necessary after its preliminary doping and dissection into slices or wafers, but before the doping of the semiconductor, before contacting, and before insertion in a case, a production step is interpolated which permits selection of semiconductors. The selection parameter is the temperature-dependence of the Hall constant, which is determined by measurement.

Document:
P 21 09 252.4

Application Date:
February 26, 1971

Disclosure Date:
September 7, 1972

Applicant:
Licentia Patent-Verwaltungs-Gmbh
6000 Frankfurt

Inventor:
K.H. Sommer

9.. REFERENCES

/136

/1/ E. Spenke — "Semiconductor Silicon" S.1-35 (Proceedings of Electrochem. Soc. Spring Meeting, 1969)

/2/ A. Herlet — Solid State Electron. 11, 717, (1968)

/3/ F.E.Gentry et al. — Semiconductor Controlled Rectifiers Prentice Hall Inc., Eaglewood Cliffs N.T. (1965)

/4/ E.F.Burtsev et al. — Sov. Phys.-Semiconductors 4, 1675 (1971)

/5/ G. Bemski — Phys. Rev. 103, 567 (1956)

/6/ W.M. Bullis — Technical Note, 571, National Bureau of Standards, Washington D.C., April 1971

/7/ S.R.Lederhandler and L.J.Giacoletto — Proc. I.R.E. 43, 477 (1955)

/8/ R.H.Kingston — Proc. I.R.E. 42, 829 (1954)

/9/ B.Lax und S.F. Neustadter — J. Appl. Phys. 25, 1148 (1954)

/10/ W. Mönch — Solid State Electron. 10, 1085 (1967)

/11/ M.Byczkowski and J.R.Madigan — J. Appl. Phys. 28, 878 (1957)

/12/ D.P. Kennedy — I.R.E. Trans. El. Dev. 9, 174 (1962)

/13/ S.K. Ghandi — The Theory and Practice of Microelectronics, Chap. 17, John Wiley and Sons Inc. N.Y. (1968)

/14/ J. Martin et al. — Solid State Electron. 9, 83 (1966)

/15/ G.J.Sprokel and J.M.Fairfield — J. Electrochem. Soc. 112, 200 (1965)

/137

/16/ J.W.Gilpin and J.C.Boatman — "Semiconductor Silicon" s. 445 (Proceedings of Electrochem. Soc Spring Meeting, 1969)

/17/ J.D.Struthers — J. Appl. Phys. 28, 516 (1957)

/18/ B.I.Boltaks et al. — Fiz. Tver. Tela 2, 2395 (1960)

/19/ W.C.Dash — J. Appl. Phys. 34, 2275 (1960)

/20/ F.C. Frank and D. Turnbull — Phys. Rev. 104, 617 (1956)

/21/ M.D.Sturge — Proc. Phys. Soc. (Lond) 73, 297 (1959)

/22/ W.R.Wilcox and T.J.La Chapelle — J. Appl. Phys. 35, 240 (1964)

/23/ D.Kendall D.B. De Vries — "Semiconductor Silicon" S. 358 (Proceedings of Electrochem.Soc Spring Meeting, 1969)

/24/ W.M.Bullis — Solid State Electron. 9, 143 (1966)

/25/ F.Heigl and R.Sizmann — published in "Crystal Lattice Defects"

/26/ H.F.John — Proc. IEEE 55, 1249 (1967)

/27/ A.J.H.de Kok — Appl. Phys. Letters 16, 100 (1970)

/28/ G.H.Schwuttke — Main Report 4. Solid State Devices Conference, Exeter, England (1970)

/29/ W.R.Wilcox, T.J.La Chapelle and D.H.Forbes — J.Electrochem. Soc. 111, 1377 (1964)

/138

/30/ J.L.Lambert — German Patent P. 21 54 373.7

/31/ H.J.Queisser — J. Appl. Phys. 32, 1776 (1961)

/32/ J.R.Carruthers et al. — J. Appl. Phys. 34, 3389 (1963)

/33/ V.N.Mordkovitch — Sov. Phys. Solid State 6, 1716 (1965)

/34/ A.Götzberger and W. Shockley — J. Appl. Phys. 31, 1821 (1960)

/35/ J.L.Lambert and M. Reese — Solid State Electron., 11, 1055 (1968)

/36/ T.Y.Tien and F.A.Hummel — J. Am. Ceram. Soc., 45, 422 (1962)

/37/ B.I.Boltaks — Diffusion in Semiconductors, Infc Earch Ltd. London (1963), Chap. 4

/38/ W. Runyan — Silicon Semiconductor Technology, Mc Graw-Hill, New York, 1965

/39/ J.L.Lambert et al. — Solid State Electron., 10, 877 (1967)

/40/ H.Reiss et al. — B. S. T. J., 35, 535 (1956)

/41/ W.Shockley and J.L.Moll — Phys. Rev. 119, 1480 (1960)

/42/ W.Shockley and J.T.Last — Phys. Rev. 107, 392 (1957)

/43/ S.F.Cagnina — J. Electrochem Soc., 116, 498 (1969)

/44/ E.D.Wolley and R. Stickler — J.Electrochem. Soc., 114, 1287 (1967) /139

/45/ M.Hansen — Constitution of Binary Alloys, Mc Graw Hill, New York, 1958

/46/ J.L.Lambert and G. Köhl — German Patent P 20 29 221.1

/47/ H.Muraoka et al. — Toshiba Review p., 49 Juli-Aug. (1968)

/48/ E.Schibli and A.G.Milnes — Mater. Sci. Engl., 2, 173 (1967)

/49/ Yu.Zibuts, L.Paritskii and S.Ryvkin — Sov.Phys.Solid State 5, 2416 (1964)

/50/ Yu.Zibuts et al. — Sov.Phys.Solid State 8, 2041 (1967)

/51/ J.S.Blakemore and C.E. Sarver — Phys. Rev. 173, 767 (1968)

/52/ C.Sah et al. — Solid State Electron.13, 759 (1970)

/53/ E.Putley — The Hall Effect and Related Phenomena London 1960

/54/ A.Hoffmann , Physical methods for impurity analysis in silicon in Halbleiterprobleme VI [Semiconductor problems VI], Fr. Vieweg u. Sohn, Braunschweig, 1961.

/55/ H.Carchano and C. Jund — Solid State Electron. 13, 83 (1970)

/56/ J.S.Moore, M.C.P.Chang and C.Penchina — J.Appl.Phys.41, 5282 (1971)

/57/ F.L.Thiel and S.K.Ghandhi — J.Appl.Phys.41, 254 (1970)

/58/ L.Rosier and C. Sah — Solid State Electron. 14, 41 (1971)

/59/ P.Doubrava — phys.stat.sol. 40, 483 (1970)

/60/ L.J.van der Pauw — Philips Res.Rep. 13, 1 (1958)

/61/ A.C.Beer — Galvanomagnetic Effects in Semiconductors Academic Press New York 1963

/62/ D.Long — Phys.Rev. 107, 672 (1957)

/63/ D.Long — Phys.Rev. 120, 2024(1960)

/64/ D.Long and J.Myers — Phys.Rev. 115, 1107(1959)

/65/ J.Messier and J.M.Flores — J. Phys. Chem. Solids 24, 1539 (1963)

/66/ H.J.Benda — Solid State Electron. 8, 189 (1965)

/67/ B.S.Alieva and V.I.Tagirov — Sov.Phys.Semic. 4, 1872 (1971)

/68/ G.W.Ludwig and R.L. Watters — Phys. Rev. 101, 1699 (1956)

/69/ E.Spenke — Elektronische Halbleiter [Electronic semiconductors], Springer Verlag, Berlin (1965)

/70/ R.J. Bassett and C.A. Hogarth — Int.J.Electron. 24, 301 (1968)

/140

/71/ E. Borchert and P. Gertkemper — Technical Report E 43/E 8, Nr.1/70 (17.2.70) AEG-TELEFUNKEN (E 43) Belecke

/72/ W.Shockley and W.T. Read, Jr. — Phys.Rev. 87, 835 (1952)

/73/ I. Somos — IEEE Transact. Communications and Electron. 83, 861 (1964)

/74/ W. Bösterling and M. Fröhlich — Elektroniker 3, 9 (1966)

/75/ J.M.Fairfield and B.V.Gokhale — Solid-State Electron. 8, 685 (1965)